ELSEVIER
ACADEMIC
PRESS

Encyclopedia of APPLIED PSYCHOLOGY



Editor-in-Chief
CHARLES SPIELBERGER

VOLUME 2 F–P1



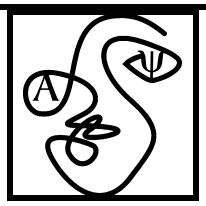
Editor–in–Chief

CHARLES D. SPIELBERGER

University of South Florida
Tampa, Florida, USA

Charles D. Spielberger is Distinguished Research Professor of Psychology and Director of the Center for Research in Behavioral Medicine and Health Psychology at the University of South Florida. Author, co-author, or editor of more than 350 professional publications, Spielberger's current research focuses on anxiety, anger, curiosity, depression, job stress, and lifestyle factors that contribute to the etiology and progression of cancer and cardiovascular disorders. His *State-Trait Anxiety Inventory* (1970, 1983), with translations in 66 languages and dialects, has become a standard international measure of anxiety. He has served as President of the International Association of Applied Psychology, the International Council of Psychologists, and the International Stress Management Association and was the 100th President of the American Psychological Association.

ЭЛЕКТРОННАЯ БИБЛИОТЕКА
Прочитал? Удали!
LIBRUS.RU



Editorial Advisors

John W. Adair
University of Manitoba
Winnipeg, Manitoba, Canada

Michael Harris Bond
Chinese University of Hong Kong
Hong Kong, China

Pieter J. D. Drenth
Vrije Universiteit Amsterdam
Amsterdam, The Netherlands

Edwin A. Fleishman
George Mason University
Potomac, Maryland, USA

Raymond D. Fowler
American Psychological Association
Washington, D.C., USA

Michael Frese
Justus-Liebig University of Giessen
Giessen, Germany

Machiko Fukuhara
Tokiwa University
Mito, Japan

James Georgas
University of Athens
Athens, Greece

Frank Heller
Tavistock Institute of Human Relations
London, United Kingdom

Stevan Hobfoll
Kent State University
Kent, Ohio, USA

Anna Leonova
Moscow State University
Moscow, Russian Federation

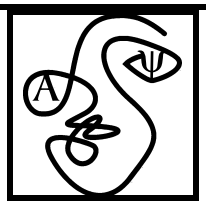
Claude Levy-Leboyer
Université René Descartes
Paris, France

Joe Matarazzo
Oregon Health Sciences University
Portland, Oregon, USA

Charles Spielberger
University of South Florida
Tampa, Florida, USA

Tuomo Tikkanen
Finnish Psychological Association
Helsinki, Finland

Bernhard Wilpert
Technische Universität Berlin
Berlin, Germany



Section Editors

Neil Anderson
Universiteit van Amsterdam
Amsterdam, The Netherlands

Andrew Baum
University of Pittsburgh
Pittsburgh, Pennsylvania, USA

John W. Berry
Queen's University
Kingston, Ontario, Canada

Mirilia Bonnes
Università degli Studi di Roma "La Sapienza"
Rome, Italy

Heliodoro Carpintero
Universidad Complutense de Madrid
Madrid, Spain

Giuseppe Carrus
Università degli Studi di Roma "La Sapienza"
Rome, Italy

Patrizia Catellani
Catholic University of Milan
Milan, Italy

Florence Denmark
Pace University
New York, New York, USA

Rocio Fernández-Ballesteros Garcia
Autónoma Universidad de Madrid
Madrid, Spain

Jane Goodman-Delahunty
University of New South Wales
Sydney, New South Wales, Australia

Yuri Hanin
Research Institute for Olympic Sports
Jyvaskyla, Finland

Ann Higgins-D'Alessandro
Fordham University
Bronx, New York, USA

Robert Hogan
Hogan Assessment Systems
Tulsa, Oklahoma, USA

Barry Kantowitz
University of Michigan
Ann Arbor, Michigan, USA

Uichol Kim
Chung-Ang University
Seoul, Korea

Howard M. Knoff
U. S. Department of Health and Human Services
Little Rock, Arkansas, USA

Frederick T. L. Leong
University of Tennessee
Knoxville, Tennessee, USA

Peter A. Lichtenberg
Wayne State University
Detroit, Michigan, USA

Ingrid Lunt
University of London
London, United Kingdom

Roy Malpass
University of Texas
El Paso, Texas, USA

Rick R. McCown
Duquesne University
Pittsburgh, Pennsylvania, USA

José M. Prieto
Universidad Complutense de Madrid
Madrid, Spain

Donald K. Routh
University of Miami
Miami, Florida, USA

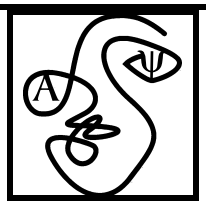
Mark L. Savickas
Northeastern Ohio Universities College of Medicine
Rootstown, Ohio, USA

Margaret Semrud-Clikeman
University of Texas
Austin, Texas, USA

Harry C. Triandis
University of Illinois, Urbana-Champaign
Champaign, Illinois, USA

W. Fred van Raaij
Universiteit van Tilburg
Tilburg, The Netherlands

Sheldon Zedeck
University of California, Berkeley
Berkeley, California, USA



Preface

During the past 50 years, there has been remarkable progress in the articulation of psychological theory, increasingly sophisticated reports of empirical research, and a dramatic expansion in professional applications of psychology in all areas of modern life. Applied psychology requires both a sound scientific foundation and attention to the practical problems that are encountered in applying psychology to an infinite variety of life settings. Growing public recognition of the potential contributions of psychology to solving life's problems has also fostered rapid growth in the number of psychologists who are engaged in research and in professional practice throughout the world.

The need for an authoritative publication that reviews and evaluates the numerous applications of professional psychology has been recognized for a number of years, and has been of special interest to the members of the International Association of Applied Psychology (IAAP), the world's oldest and largest membership organization of individual psychologists. The interests and concerns of IAAP members, who reside in more than 80 countries, have been strongly noted by former IAAP presidents Edwin A. Fleishman, Claude Levy-Leboyer, and Harry C. Triandis.

Significant contributions to international applications of psychology were explicitly recognized and discussed in the context of their diverse societal settings by Bernhard Wilpert in his Presidential address at the IAAP's 24th International Congress of Applied Psychology in 1998. It was my honor and pleasure to succeed Professor Wilpert as President of the IAAP at this Congress, which was held in San Francisco, California, and to begin participating in discussions to identify the most important contributions to applied psychology that are reported in this encyclopedia.

The development and current status of applied psychology can be best understood by examining the historical evolution of the field, which is cogently described by Helio Carpintero, the official Historian of the IAAP. In his overview article on the history of applied psychology, Carpintero highlights the benefits resulting from psychologists joining together to form professional associations that have facilitated the development and integration of psychological theory and research and that have stimulated and guided more effective applications of psychology in professional work.

Founded in 1920 as the *Association Internationale des Conferences de Psychotechnique* (International Association for Psychotechnology), the IAAP adopted its present name in 1955. Edouard Claprēde convened and chaired the first meeting of the association, held in Geneva, Switzerland in 1920, and served as its President for more than 20 years. Meetings were held in a number of European cities, including Barcelona, Milan, Paris, Moscow, and Prague. An excellent history of the evolution and development of applied psychology and the IAAP can be found in the address given by former President Edwin A. Fleishman on receiving the 1999 American Psychological Association's Award for Distinguished Contributions to the International Advancement of Psychology (*American Psychologist*, November, 1999, pp. 1008–1016).

For recommending that the IAAP take an active role in sponsoring and developing the *Encyclopedia of Applied Psychology*, we are indebted to George Zimmar, an active IAAP member and former senior sponsoring editor at Academic Press, now Elsevier. The first step in developing the encyclopedia was the appointment of an Advisory Board composed of distinguished colleagues representing all areas of

applied psychology. In addition to recommending the major areas and the specific topics to be covered in the encyclopedia, the Advisory Board identified highly qualified colleagues who would be invited to serve as Section Editors for the selected areas.

The responsibilities of the Section Editors included writing a major article to provide an overview and analysis of their respective areas and working with the Editor-in-Chief and the Advisory Board in identifying the specific topics to be included in each area. They also helped to determine the amount of space to be allocated for each topic and recommended colleagues with outstanding knowledge and expertise to serve as the authors for the individual articles. In addition, the Section Editors reviewed and edited the articles that presented information on the topics that were related to the subject matter that was included in their sections.

The articles that are included in the encyclopedia are presented in alphabetical order, as determined by the title of each article. Consequently, the underlying section organization will not be readily apparent. At the conclusion of each article, the reader is directed to additional articles that include related information. The authors of each article also provide a number of references that are recommended for further reading. It should be noted, however, that it was not possible to cover all of the many areas of applied psychology. Several emerging areas are not included because we failed to recognize their importance at the time the topics were selected for the encyclopedia. It is also unfortunate that a few authors were unable to meet the deadlines for submitting the manuscripts for their articles.

The *Encyclopedia of Applied Psychology* is truly a product of the collaborative efforts of the members of the IAAP. We are especially indebted to former IAAP Presidents Edwin A. Fleishman, Harry C. Triandis, and Bernhard Wilpert, and to our current president, Michael Frese, for their leadership, encouragement, and tangible contributions to the encyclopedia. For their dedicated contributions to the encyclopedia, I would also like to thank the members of the Advisory Board and the Section Editors, whose names are listed in the front matter. Finally, on behalf of the IAAP, I thank the authors who contributed articles in their areas of expertise and the editorial and production staff of Elsevier for the tremendous amount of time and energy that was needed to bring this large project to fruition.

CHARLES D. SPIELBERGER, Ph.D., ABPP
President, IAAP, 1998–2002

Fair Treatment and Discrimination in Sport

Vikki Krane

Bowling Green State University, Bowling Green, Ohio, United States

1. Introduction
2. Social Prejudice in Sport
3. Celebrating Difference through Sport
4. Social Equity in Sport
5. Applied Sport Psychology and Social Justice
Further Reading

GLOSSARY

classism Differential treatment based on social class or economic status; the privileging of the wealthy social group so that that group has greater power and status than other social class groups.

equality Identical treatment across social groups.

equity Fair treatment, but not identical treatment, across social groups; equal access to necessary opportunities and resources.

ethnicity A social category characterized by cultural heritage, including norms, values, and customs.

ethnocentrism Differential treatment based on ethnicity; the privileging of one ethnic group so that the group has greater power and status than other ethnic groups.

heterosexism Differential treatment based on sexual orientation that privileges heterosexuals over other sexual orientations; stigmatizes behaviors, identities, and relationships that are not heterosexual.

nationalism Differential treatment based on national origin or country of residence; the privileging of people of one nation, or group of nations, so that the group has greater power and status than people of other nations.

race A social category characterized by biological or genetically transmitted traits.

racism Differential treatment based on race; the privileging of one racial group so that the group has greater power and status than other racial groups.

sexism Differential treatment based on gender; the privileging of one gender so that the group has greater power and status than other ethnic groups.

social class A social category based on economic position in society, including income, wealth, education, occupation, and social connections.

social justice Social tolerance ensures that people of all social groups have the necessary resources and opportunities.

social prejudice Systematic mistreatment, bias, or bigotry toward people based on social group membership.

stereotypes Characteristics perceived to be shared by members of a social group, without consideration for individual differences.

Fair treatment in sport occurs when there are equitable resources and opportunities for all participants, regardless of social group membership (e.g., gender, race, ethnicity, nationality, sexual orientation, and social class). Sport offers the opportunity to traverse social identities, but it also can be divisive. In some cases, differential treatment occurs in which minority social group members are treated unfairly. In other

©2004 Elsevier Inc.
All rights reserved.

situations, sport is an avenue to educate people and increase awareness about a wide range of social issues, fight social injustice, and provide humanitarian assistance. Applied sport psychologists may employ a variety of strategies to promote fair and equitable sport.

1. INTRODUCTION

Sport often is considered a microcosm of society: Common practices in everyday life are replicated in sport environments. Issues such as equitable or non-equitable conduct found in society generally are reproduced in sport settings. Sometimes sport seems to traverse racial, ethnic, social, national, and gender boundaries; sport has the power to bring diverse people together while deemphasizing social or cultural differences. International sport is predicated on the ideal of achieving peace and friendship among world peoples. It provides a common language, connects people worldwide, and offers an avenue for cultural exchange and enhanced understanding across national boundaries. This is seen in events such as the Olympics, Football World Cup, Tour de France, and the Commonwealth Games. Athletes who may be from divergent cultural groups will cooperatively engage in sport. This has occurred with such disparate groups as Iranians and Iraqis or Arabs and Israelis. In fact, competitions have been created specifically to mend political strife. The Goodwill Games, for instance, were constructed to bring together the United States and former Soviet Block nations after the boycotts of the 1980 and 1984 Olympic Games.

Although sport offers the opportunity to traverse cultural identities and unite different peoples, it also can have the opposite effect. Cultural differences may be exacerbated through sporting events, which in turn become propaganda for the victorious side. Sport reproduces dominant ideologies and is accessible disproportionately to members of privileged social groups. Obviously, these matters negatively affect sport experiences.

2. SOCIAL PREJUDICE IN SPORT

People naturally group like individuals into social groups; this categorization is a perceptual process that allows individuals to make sense of the vast social world. Stereotypes emerge as people identify characteristics perceived as common among the members of a social group. Sometimes these common characteristics are benign (e.g., basketball players are tall). However, some stereotypes may be incorrect (e.g., black people cannot swim) or even malicious (e.g., Islamic men are violent or terrorists). Prejudice occurs when stereotypes lead people to have negative attitudes toward all members of certain social groups, whereas discrimination occurs when people behave negatively toward members of a social group.

Unfair treatment in sport, as in society at large, often arises from prejudices such as racism, ethnocentrism, sexism, heterosexism, classism, and nationalism. These terms reflect various forms of social exclusion based on socially constructed categories or identities (e.g., gender, race, nationality, sexual orientation, and ethnicity). Identities are socially constructed as social norms and expectations create and reinforce attitudes toward individuals based on social categories. For example, beliefs about how males and females should behave and look are based on social customs. These customs create expectations for males and females that affect actual behavior, values, and beliefs. Social beliefs and social norms often lead to stereotypes about certain groups of people, which reflect societal values and convey acceptable and unacceptable behaviors and values. They also deprive people of their individuality and limit opportunities.

Stereotypes and negative attitudes are the major roadblocks to sporting achievement, not lack of ability. Innately, there is no reason that people in various social groups should have any less athletic talent than individuals in majority social groups, but examples of disproportionate accomplishments abound. For example, far fewer African, Latino/a, or Asian athletes are represented on Canadian, UK, or Australian professional sport teams compared to white athletes. Stereotypes such as "black players are not as smart as white players" and "Asians lack athletic ability" impede their progress in sport. Beliefs that "women are not interested in sport" or that "Aboriginal people are not as 'civilized' or cognitively developed as white people" reduce their opportunities in mainstream (i.e., majority-ruled) sport.

Because these athletes are not expected to succeed, it is likely that coaches will treat them differently than athletes expected to be more skilled. The athletes for whom coaches have low expectations likely receive less constructive feedback, less praise, fewer practice opportunities, fewer chances to correct mistakes, and more criticism. This feedback, in turn, may decrease confidence and increase anxiety or may not provide learning opportunities for certain groups of athletes. Thus, these athletes do not attain the same skill level as

athletes from majority social groups, reinforcing the correctness of the stereotype. The concomitant lack of successful role models and inability or unwillingness of sport clubs to support minority athletes further minimize their chances of excelling in sport. Additionally, individuals with minority identities are often socially or economically disadvantaged and therefore do not have the resources to develop their athletic abilities.

Cultural and prejudicial attitudes evolve and change over time; therefore, they are socially and historically bound. For example, although today black players appear to be represented proportionately in British football, Asian players are not. However, not long ago, black players were struggling for representation. As social values change, so do attitudes toward various social groups in sport. Judgments about specific social groups also may differ across cultures. French fans, for instance, backed openly lesbian, professional tennis player Amelie Mauresmo, whereas in other countries she would not be supported. There is not a single way to explain prejudice or inequity in sport because it is played out in a myriad of ways given the dominant culture where it occurs. The underlying commonality, though, is that opportunity, status, and power are disproportionately afforded consistent with the social hierarchy supporting the dominant social group.

2.1. Racism

Inequitable treatment due to racial or ethnic identities occurs across sporting environments. Racism affects university, professional, and elite sports, from youth to masters' levels. It has been documented in sport worldwide. Incidents of racist fan behavior at World Cup or international rugby events, for example, are not uncommon. Racist actions in sport result in fewer racial minority athletes engaging in particular sports (e.g., swimming and golf), minimal numbers of athletes of color competing at elite levels of sport, few individuals of color as coaches or managers of athletic programs, overt prejudicial behaviors toward racial minority athletes, lack of fan support for minority athletes, and limited opportunities to learn sports. Conflated with these indignities, female athletes of color face additional barriers, as do disabled, indigenous, or lesbian, gay male, bisexual, or transgendered (LGBT) athletes of color.

A common racist action in sport is stacking, which occurs when a certain social group of athletes is funneled into peripheral positions on athletic teams. For example, in Western team sports, white athletes historically held the leadership and central positions. (White) coaches believed that black athletes were not smart enough to take on the responsibility of this position. Rather, black athletes were guided into positions requiring speed and jumping ability (e.g., pass receivers). This practice led coaches to encourage young black players to gain skills in peripheral positions if they wanted an opportunity to participate in higher levels of the sport. It also encouraged self-selection into positions where there were black role models. Stacking ultimately limited the number of athletes of color on a team because these players competed for only a few positions. This practice occurs across a wide array of sports.

Although the overt practice of stacking has declined in many sports, lack of minority representation in sport administration and coaching is widespread. Because these are the powerful, decision-making positions, minority athletes have little representation in how sport is conducted or who is selected for teams or management positions. Additionally, sport has become a big business. The owners, managers, and organizers in the business of sport have become quite wealthy and powerful, allowing them to wield considerable power concerning how sport is run and who can be involved. Few men of color rise to these positions and almost no minority females are sport team owners or managers.

2.2. Sexism

Research on gender in sport consistently documents fewer opportunities and less support for female sport participants compared to males. International competitions, such as the Olympics, have more events for male athletes; governing bodies are male dominated so women have little voice in how sport is managed; the media reinforces gendered ideologies and provides limited coverage of female athletes; sexual harassment toward female athletes is widespread; and females in sport must contend with stereotyping and "accusations" of being lesbians. These actions have led to sport being considered a patriarchal institution; that is, it is dominated by males who hold social, political, and economic power.

Sport also reinforces and supports hegemonic masculinity and femininity. These ideals reinforce socially sanctioned mannerisms specific for males or females. Most societies encourage males to be masculine and females to be feminine. Characteristics associated with hegemonic masculinity include strength, competitiveness, assertiveness, confidence, and independence. Being gentle, compassionate, emotional, and dependent, for example, are

characteristics of hegemonic femininity. These characteristics are considered hegemonic because they are accepted and reinforced by social norms, customs, and behaviors. Hegemonic customs and behaviors are so widely visible that they are rarely questioned; rather, they are accepted as "natural." Males who act contrary to hegemonic masculinity or females who counter hegemonic femininity face discrimination. Because sport is considered a masculine domain, especially contact sports, there is more support (financially and socially) for boys and men in sport than for girls and women.

Fundamental, non-Western societies are even more restrictive toward women's sport participation. Within traditional Islamic cultures, women face rigid barriers for participation in sport. This culture also emphasizes modesty, and for women this encompasses completely covering their bodies in clothes. This expected attire may impede participation in sport altogether or limit women's ability to compete in elite events in which competitive attire is inconsistent with Islamic culture. Only in all-female environments may Islamic women wear shorts and t-shirts; otherwise, they are expected to be covered properly. For Islamic women to compete at elite events, they must forsake traditional attire and modesty, participate in mixed-gender venues, and resist the strongly patriarchal social structure. They also risk severe reprisals for doing so. Female Islamic athletes have received public disavowals and death threats for participating in settings not consistent with Islamic law. Although facing different ideologies of intolerance, Aboriginal women have similar struggles. Whereas women in general are underrepresented in sport, women who are nonwhite, economically disadvantaged, older, physically challenged, or lesbian, bisexual, and transgendered are even less visible.

2.3. Heterosexism

LGBT individuals in sport face paradoxical circumstances. On the one hand, they may be able to "pass" and be undetected in sporting environments. In which case, they may avoid overt discriminatory actions based on their sexual orientation. On the other hand, that they must conceal their LGBT identities suggests the depth of prejudice toward this social group. LGBT athletes often are perceived as countering hegemonic femininity or masculinity, which appears to underscore prejudice against this social group.

Bias against LGBT athletes is well documented in Western societies. They face verbal harassment, lack of media attention and endorsements, negative bias by officials or judges, and prejudicial treatment by administrators and coaches. Incidents of physical and verbal abuse by teammates are also documented. Some sport organizations have rules specifically aimed toward prohibiting LGBT participants. For example, the U.S. Ladies Professional Golf Association only allows women who were born female to compete, effectively barring transsexuals from participation. An athlete marching in the opening ceremonies at the 1998 Gay Games exemplified the heightened fear of detection among LGBT athletes from fundamental Eastern countries. Holding a placard for Lebanon, he wore a mask over his face. In Lebanon and other countries with fundamental religious-based governments, LGBT individuals may be imprisoned because of their sexual orientation.

2.4. Nationalism

As international competitions bring together individuals from around the world, they also encourage nationalism. International sport and politics seem inextricably linked. Often, governmental involvement in sport reflects a desire for global recognition or prestige. In the Olympics, for example, if the ideal of promoting world togetherness was truly embraced, news agencies would not report the numbers of medals won by participating countries. However, these medal counts are used by various countries to confirm their sovereignty. This was most evident during the Cold War era when the campaign for democracy versus socialism was partially played out in the international sport scene.

International sport is important economically in terms of world recognition and global esteem. Today, rising industrial nations vie to host the Olympics, and other prominent international sport events, to demonstrate their new strength in the international market. Nationalism also is seen as the sport media constructs national identities through coverage of major sporting events. Positive national identities are reinforced, whereas stereotypes about, and difference between, athletes from other countries are accentuated. Nationalism has led to situations of poor sporting behaviors, such as fan hooliganism, fights during competitions, and unfair judging or refereeing. Additionally, some national sport organizations, such as in Japan, only support people of local heritage. Foreigners and immigrants are excluded due to institutionalized nationalism.

2.5. Religious Intolerance

As previously discussed, religious beliefs may influence sport experiences. One social group particularly affected by religious influence on sport is Islamic women. Discrimination based on religion may include situations in which important sporting events are scheduled on non-Christian religious holidays, religious expectations are not accommodated, religious customs are belittled, prejudicial jokes are shared or prejudicial comics are displayed, or players are overtly discriminated against because of their religious beliefs.

Perhaps the most evident occurrence of religious intolerance manifest in sport was at the 1972 Olympics when 11 Jewish athletes and coaches were assassinated by terrorists. In the current global climate, sportsmen of Arab dissent are viewed cautiously and considered potential terrorists due to their heritage. Also, the revealing uniforms common in Western women's sport are not consistent with the modesty espoused in many Asian and Islamic cultures. Still, some international sport organizations require athletes to wear these styles. Similarly, many sport settings do not accommodate the religious practices of Islamic athletes who engage in obligatory prayer times.

2.6. Interactions among Social Prejudices

The various forms of social exclusion interact to create a hierarchy of status and power in sport. For example, when examining racism in sport, the focus is often on men's sport. The omission of considering race in women's sport is a form of sexism and neglects consideration of the interaction between racism and sexism. Female athletes of color face challenges related to their gender as well as their race. Across sporting cultures, males seem to reap the most privilege, particularly white, heterosexual males. Male athletes of color are afforded the privilege of being male, although they are marginalized due to their race or ethnicity. In countries where people of color are the high-status majority, heterosexual males still garner the most privilege and power. Additionally, it is important to consider the influence of social class. When examining sport across diverse countries, people of color or ethnic minorities disproportionately are the most socially disadvantaged people. Females are also represented disproportionately in socially disadvantaged classes; in this case, racial and ethnic minority females face the largest social and economic challenges. Throughout the world, LGBT individuals face much discrimination, regardless of, or in addition to, their race, ethnicity, social class, or gender. However, white, gay males still have more social benefits than females or people of color in these communities. All of these patterns of gender, racial, ethnic, and sexual orientation superiority are consistently evident across Western societies. In non-Western societies, there may be different cultural forces affecting different patterns of privilege and discrimination, but discernable patterns are evident. As global sport becomes more commodified, minority participants continue to become even more distanced.

3. CELEBRATING DIFFERENCE THROUGH SPORT

A number of sporting events have been developed to provide safe and comfortable environments for individuals who commonly experience discrimination in traditional sporting venues. These events cater to specific social groups and include competitive, social, and cultural activities. Examples include competitions for Jewish athletes, LGBT athletes, Arab athletes, and Islamic athletes. Such events promote cultural pride and self-esteem. Participants engage in competitive events, but they also are surrounded by individuals with similar social identities. They are able to work toward sporting goals and immerse themselves in a setting that celebrates their unique culture. Athletes are freed from discrimination and inequity, creating a personally and athletically empowering environment.

The Maccabiah Games are an Olympic-style event held in Israel every 4 years. The best Jewish athletes in the world are invited to participate in athletic, cultural, and educational events. The games offer the opportunity to share participants' customs and heritage in a competitive setting. The Gay Games offer a similar setting for LGBT athletes. Every 4 years the Gay Games celebrate the sporting and cultural practices of the participants. These games began as a venue for competition in an environment free from discrimination. They also counter stereotypes about LGBTs as well as celebrate the distinctiveness of LGBT life.

The Arab Games were created as a competitive event for male athletes from the nations of the Arab League. These games cater only to Arab males; participation by women is contrary to dominant religious beliefs in these countries. Recently developed, the Islamic Countries Women's Sport Solidarity Games provide a

competitive environment for women consistent with Islamic law. Competitive events are held in an all-female environment. Males may be spectators only at the opening ceremonies, where the women wear traditional Islamic attire.

To promote national pride and enthusiasm in European small nations, the Games of the Small States of Europe bring together athletes from Andorra, Cyprus, Iceland, Liechtenstein, Luxembourg, Monaco, San Marino, and Malta. Because athletes from these nations generally cannot keep up competitively with athletes from larger countries, these games provide the opportunity to compete successfully at an elite level. The North American Indigenous Games celebrate native people in the United States and Canada. A variety of competitions also exist for Australian indigenous peoples. These events often include traditional Aboriginal games as well as contemporary sport competitions.

The personal and cultural influence of these types of competitions is remarkable. For many athletes, these are the only places where they can compete devoid of prejudice and discrimination. Personal, cultural, and athletic empowerment often result. Athletes celebrate their shared history and culture while also engaging in international competition. Typically, there is extraordinary energy and excitement in such compassionate environments. However, the necessity for these types of competition may be questioned because they may be perceived as separatist. In the current social climate, worldwide, these competitions seem to provide an important experience for the athletes. When athletes with devalued social identities can participate comfortably and safely in integrated competitions, perhaps independent competitions no longer will be needed.

4. SOCIAL EQUITY IN SPORT

Although examples of inequitable treatment in sport are evident, in many ways, sport brings together people from diverse social groups. It has the potential to counter stereotypes and inequitable treatment of minority athletes. For example, Amir Hadad, an Israeli Jew, and Aisam-Ul-Haq Qureshi, a Pakistani Islamic, competed as doubles partners in the 2003 Wimbledon tennis tournament. During international events such as the Olympics, diverse athletes are seen interacting on the playing fields as well as socializing in the Olympic Village. After the September 11, 2001, terrorist attack on the United States, American athletes were met with great encouragement and support throughout the world.

Sport has become a platform to increase awareness of health concerns such as breast cancer and AIDS. Races and bicycle rides are used to raise money for research and assistance programs. Examples of sport as social agent are widespread: In Kenya, sport helped fight AIDS; through sporting events money was raised to fight cancer in the United Kingdom; sport programs are used to reduce crime in Germany and the United States; and sports assist in overcoming religious intolerance in Ireland and the Middle East. Many sport stars have created programs to reach out to disadvantaged children to provide them the benefits of sport. For example, Tiger Woods created the Tiger Woods Foundation, which introduces golf to minority and at-risk, inner-city youth. Sport organizations, such as the Professional Golf Association and the National Hockey League, are also reaching out to minority youth. The potential of sport to educate people and increase awareness about a wide range of issues, fight social injustice, and provide humanitarian assistance is immeasurable.

4.1. Opposing Social Prejudice through Sport

Because of the wealth, status, power, and popularity associated with sport, it can be a strong catalyst for social change. Many national and international sport organizations are working toward more equitable sport environments. A wide array of international sport organizations and conferences have addressed issues of inequity in sport. Sporting organizations are making policy to minimize sexism, racism, and heterosexism. Attempts to provide sport opportunities for economically disadvantaged people are also incorporated into these policies.

4.2. Reducing Racism in Sport

Perhaps the most poignant example of sport as an agent of social change was during the anti-apartheid movement. Most international sport events banned South Africa from participating until apartheid was abolished. Along with other political pressures, South Africa eventually succumbed and changed its social system. This international campaign in sport helped bring down the horrific racism in South Africa and permanently changed international sport. This effort toward reducing racism continues today. For example, participation in the Commonwealth Games is contingent upon commitment to countering racism in disadvantaged countries.

Sport, the Olympic Games in particular, led to serious reconsideration of the inequitable treatment of Aboriginal people in Australia. In response to national protests after being awarded the 2000 Olympic Games, the International Conference on Human Rights and Sports was held in Sydney in 1999. Here, the plight of the Aboriginal people was highlighted and propositions to reduce prejudice and discrimination were developed. The Indigenous Sport Program, developed to increase sport opportunities at all levels as well as provide economic assistance for rising elite Aboriginal athletes, was one outcome of this conference.

Many professional sport organizations have enacted policies specifically to counter racism. FIFA, soccer's world governing body, held an antiracism conference in Buenos Aires in 2001. Delegates from 200 football (soccer) national associations engaged in brainstorming to develop strategies to minimize racism in football. Ultimately, FIFA passed a resolution expecting that all clubs, associations, and officials actively take steps to eradicate racism in football. In 1999, Football against Racism in Europe (FARE) was developed, involving organizations from 13 countries. FARE assists and links various local and national endeavors to counter racism in football (soccer). By 2003, all 92 European professional soccer clubs supported the "Kick Racism out of Football" campaign. "Show Racism the Red Card" is another antiracist education campaign in the United Kingdom. This program engages professional footballers as antiracist role models. Sportscotland developed a comprehensive equity strategy for sport. It encourages all sport governing bodies to address inequities related to race, ethnicity, gender, sexuality, ageism, religion, and disability. Similar programs have emerged throughout the world.

4.3. Reducing Sexism in Sport

A variety of international organizations to counter sexism in sport have emerged. The International Association of Physical Education and Sport for Girls and Women was founded in 1949. This group represents and lobbies for women and girls in sport and physical education at the international level. It is a member of the International Council for Sport Sciences and Physical Education and therefore is involved in developing international policy regarding women's sport.

In 1994, the International Working Group on Women and Sport was established during the first World Conference on Women and Sport. This group drafted the Brighton Declaration, which implored governments and sport organizations to "comply with the equality provisions of the Charter of the UN, the Universal Declaration of Human Rights, and the UN Convention on the Elimination of all Forms of Discrimination Against Women." The Second World Conference on Women and Sport was held in Windhoek, Namibia, in 1998. It resulted in the Windhoek Call for Action, which reaffirmed these principles. In response to these declarations supporting women's sport, many conferences and resolutions worldwide followed. For example, in 2001, the First Asian Conference on Women and Sport was held in Osaka, Japan. At this conference, delegates reaffirmed their commitment to the Brighton Declaration and the Windhoek Call for Action and dedicated themselves to the advancement of women in sport.

In addition, the Women's Sport Foundation in the United States, United Kingdom, and Australia lobbies for increased sport opportunities for women and girls in sport. In Canada, the Canadian Association for the Advancement of Women in Sport works with Sport Canada to increase opportunities for girls and women in sport and to ensure healthy environments for female participants. Similarly, the Australian Sports Commission expressly supports women's sport. The European Women and Sport organization, representing 41 nations, promotes gender equality in European sport organizations.

National policies to reduce sexism and increase opportunities for women have been applied to sport. In the United States, Title IX makes it illegal to discriminate against individuals based on sex in any public educational setting, including sport. This law greatly affected women's sport by decreeing that girls and women receive equitable opportunities, funding, coaching, and facilities. Although still not equal to boys and men, since the enactment of Title IX, the number of girls and women participating in sport, and social acceptance of female athletics, has greatly increased. In Canada, the Charter for Rights and Freedoms protects the sport opportunities for girls and women.

4.4. Securing Child Welfare in Sport

Concern for the well-being of child participants is another area in which sport organizations encourage fair treatment. An international organization, Defense for Children International, based in Switzerland, raised concerns about the exploitation of children in elite sport. It advocated that children should have the following rights in sport: "(i) to do as one chooses with one's body, (ii) to be consulted, (iii) to physical and

psychological integrity, (iv) to associate freely with whom one chooses, and (v) to education." In the United Kingdom, the British Association for Sport and Exercise ratified a welfare policy mandating that its members safeguard and protect children and vulnerable adults from sexual abuse.

4.5. Economic Reforms through Sport

Through sport, international awareness has increased concerning the horrendous working conditions of sweatshops in developing countries. United Students against Sweatshops (USAS), an international student movement among university students, has led the campaign fighting for better working conditions and rights for workers. It has lobbied universities and colleges to be proactive and discerning when making contracts with various sporting goods companies that sell items with university logos. Groups such as USAS have led boycotts against sporting goods companies (e.g., Adidas, Nike, and Umbro), imploring the elimination of exploitive child labor, procurement of fair wages, and safe work environments.

Although all the organizations and sport group actions described previously focus on specific social justice issues (e.g., racism or sexism), they also recognize the interactions among these issues. For example, groups fighting sexism also forward policies against heterosexism and racism. While focusing on the abysmal working conditions in sweatshops, it is apparent that women and children of color are most likely to be engaged in these positions. Sport organizations such as national Olympic committees, professional sport organizations, and national governing bodies have programs encouraging diverse athletes to participate in sport as well as policies countering sexism, racism, and other social injustices.

When considering fair play in sport, there are two sides to the issue. Certainly, inequitable treatment of diverse sports people occurs. Individuals with minority social identities may be marginalized and devalued in sport; they may face stereotypes, prejudice, and barriers. However, sport also brings together diverse people. In the name of fair play, sport celebrates achievements of all people and provides an environment in which divergent social groups can leave behind political agendas and simply engage in sport. Sport also has become a tool for social justice activities. Through sport, intolerance is challenged and social change continues.

5. APPLIED SPORT PSYCHOLOGY AND SOCIAL JUSTICE

Given the potential of sport to encourage social justice and to challenge unfair treatment, sport psychologists are in an ideal position to effect such action. Applied sport psychologists work with athletes, coaches, parents, and administrators to develop productive sport behaviors leading to improved or consistent performance and enjoyment. Inequitable sport settings limit both of these possible outcomes. Thus, efforts to resist inequitable treatment logically complement more customary sport psychology goals.

Challenges to inequitable treatment may be focused on a societal level, as previously discussed, or it may be approached on a local level by creating positive and fair team environments. Teaching young athletes to value fairness and rebuke prejudice and discrimination will eventually affect all of sport. One team or league can make a major difference. For example, when teenage members of a golf team refuse to compete because a teammate is not allowed to play due to his or her skin color, their action speaks volumes to the larger society. Members of sport teams that engage in community service learn about people living in different situations than their own. College athletes in the United States, for instance, help build shelters for homeless people and tutor economically challenged children. Sport psychologists can engage the team in discussions highlighting the social justice issues inherent in such settings.

Many sport psychologists already endorse the importance of developing constructive team climates through creating mastery or task-oriented sport environments. In such settings, coaches emphasize skill development, cooperation, improvement, and development of peer relationships. In contrast, performance- or ego-oriented climates emerge when coaches encourage interpersonal competition and promote a "win at all costs" attitude. When athletes are rewarded for effort, improvement, and teamwork, they are likely to have a long-term investment in sport, use healthy training techniques, and adopt positive sporting behaviors. These athletes also tend to believe that success in sport is due to hard work and working cooperatively with others. Athletes who are rewarded primarily based on social comparison (e.g., winning) are likely to be willing to do whatever it takes to win, such as using aggressive or harmful behaviors including trash talking in a prejudicial manner or other overtly discriminatory behaviors.

A logical extension of traditional means of creating a task-oriented team setting is to address issues of social justice and encourage fair treatment of all athletes. It seems reasonable that athletes who participate in a task-oriented climate will be likely to accept the importance of being sensitive to individual differences. If, for example, athletes of different races or ethnicities are dedicated to team improvement and are invested in being successful, they will also be willing to work together in a productive manner. Creation of a task-focused environment may be the first step toward promoting fair treatment within sport teams or sport leagues.

Sport psychologists and coaches wanting to challenge unfair treatment due to prejudice may employ a variety of strategies. An initial, but important, step in this direction is to consider the language used by leaders and team members. Coaches and sport psychologists should model appropriate language. Racist, sexist, homophobic, or any other prejudicial comments should not be tolerated. As individuals point out the inappropriateness of such language and reinforce its unacceptability, athletes will become more aware of how their words may be hurtful. Situations in which prejudice or discrimination occur provide opportunities to open dialogue about sensitive issues. Furthermore, most strategies to encourage team cohesion will also promote acceptance of diverse people within a team. Creating team goals that include cooperative efforts, providing athletes opportunities to learn about athletes with different social identities, and emphasizing team diversity as a strength will lead to enhanced cohesion while also minimizing prejudice. Athletes should also be encouraged to "do the right thing." Empowering athletes to acknowledge and challenge prejudice and discrimination will further increase the likelihood that sport settings will promote fair treatment of all participants.

Ensuring equitable treatment in sport will require the involvement of athletes, coaches, and administrators. Sport psychologists are positioned to interact and educate each of these factions. As sport psychologists make social justice in sport a priority, it is more likely that sport will achieve its potential to bring together diverse people for fair competition, create avenues for cultural exchange, and ensure positive experiences for majority and minority athletes alike.

See Also the Following Articles

Diverse Cultures, Dealing with Children and Families from ■ Gender and Culture ■ Homophobia ■ Homosexuality ■ Prejudice and Discrimination ■ Sexual Behavior ■ Sexual Behavior and Culture ■ Stereotypes

Further Reading

Andersen, M. (Ed.) (in press). *Understanding applied sport psychology*. Champaign, IL: Human Kinetics.

Carrington, B., & McDonald, I. (Eds.) (2002). *"Race," sport and British society*. London: Routledge.

Coakley, J., & Dunning, E. (Eds.) (2000). *Handbook of sports studies*. London: Sage.

Collins, M. F. (2003). *Sport and social exclusion*. London: Routledge.

Griffin, P. (1998). *Strong women, deep closets: Lesbians and homophobia in women's sport*. Champaign, IL: Human Kinetics.

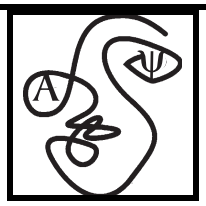
Hargreaves, J. (2000). *Heroines of sport: The politics of difference and identity*. New York: Routledge.

Kidd, B., & Donnelly, P. (2000). Human rights in sport. *International Review for the Sociology of Sport*, *35*, 131–148.

Maguire, J. (1999). *Global sport: Identities, societies, civilizations*. Malden, MA: Blackwell.

Messner, M. A. (2002). *Taking the field: Women, men, and sports*. Minneapolis: University of Minnesota Press.

Schaffer, K., & Smith, S. (Eds.) (2000). *The Olympics at the millennium: Power, politics, and the Olympic Games*. Piscataway, NJ: Rutgers University Press.



Family and Culture

James Georgas
University of Athens, Athens, Greece

1. Culture
2. Family Theories
3. Family
4. Kinship Patterns and Family Matters
5. Family Roles and Power
6. Psychological Dimensions
7. Conclusion
   Further Reading

There are two general approaches to the study of family. One approach has been the focus on the family, its structure, interactions between family members, child development, and other issues. The second approach has been the study of the family within its cultural context. This approach assumes that the family is one system within other social systems of a culture, and that the study of the family necessarily entails studying its relationships with the institutions and other aspects of the culture. In addition, the usual approach in cultural psychology studies the family in one culture or society, whereas the cross-cultural approach is comparative, studying universals and differences in families across cultures or countries. Cultural anthropology, sociology, psychology, cross-cultural psychology, psychiatry, economics, political science, and geography have each contributed to the study of family from the perspective of its own discipline. However, there is considerable overlap in theories and methods so that one can also talk about a multidisciplinary approach employed by many in each field. A number of issues are related to the study of family and culture. There are different family types: nuclear families, extended families, and one-parent families. Family types differ according to subsistence patterns, social systems, religion, and other cultural features. Marriage takes place in different ways in different cultures. Kinship ties, postmarital residence, and descent are important determinants of family types. The social power of the mother and the father is related to the institutional structure and values of the culture. Socialization of children by the family is related to ecocultural variables and family types and results in different psychological characteristics across cultures.

1. CULTURE

Culture is a very broad term referring to all human-made patterns of activity, including societal institutions, practices, norms, rules, means of subsistence, symbols, communication, settlements, transportation, art, religious beliefs, ideas about nature and humans, interactions among humans, child-rearing practices, values, beliefs, cognitive processes, personality, and social behavior. There are many definitions of culture, each stressing different elements, including the patterns of human activity, the unique symbols in each culture, the symbols transmitted through language, the achievements of the group, the artifacts, the transmission of the traditions from generation to generation, and the mutual process of humans creating cultural elements and culture-shaping human and social patterns.

©2004 Elsevier Inc.
All rights reserved.

The term culture, however, is also used to designate a group with a common heritage as described in the previous definition of culture. That is, culture is used as a synonym for "ethnic group," "nation," "society," or "country." Anthropologists initially employed the term to refer to small exotic groups of people, often referred to in the past as primitive peoples, savages, or indigenous groups. Employed in this sense, larger geopolitical groups are referred to as nations or countries. However, culture is also employed in taxonomies of nations, such as Western culture or Islamic culture. If culture is used to define an ethnic group with a specific "culture," then the number of cultures has been estimated by anthropologists to be in the thousands, whereas there are approximately 200 nations or countries throughout the world.

Thus, culture is employed in both ways in family research: It refers to a specific country or to those elements that characterize the society.

2. FAMILY THEORIES

Family theories have been developed in sociology, cultural anthropology, psychology, psychiatry, and cross-cultural psychology. The first family theories were developed in the 19th century. A form of social Darwinism theorized that the structure and function of family adapted, as a social organism, to the environment. Marxist theory employed the concept of economic determinism to explain how economic resources determined social power, which in turn determined class struggle. The concept of social power was extended to explain authority and power structures of the father and mother. Durkheim's theory explained the existence and the changes in family structure and function in terms of the family's functional role in the preservation of society. Functionalism analyzed the role of family as part of a greater whole, in which an equilibrium with other social institutions was established. This was essentially a systems theory, in contrast to the biological or economic determinism of social Darwinism and Marxist theory.

One of the most influential theories of family in the second half of the 20th century was Parsons' structural–functional perspective, in which society was viewed as an organism striving to resist change and to maintain itself in a state of equilibrium. Structure refers to the members of the family (parents, child, and kin), and function refers to how families satisfy physical and psychological needs for survival and maintenance. According to Parsons, the adaptation of the family unit to the industrial revolution required a nuclear family structure that could carry out societal functions and satisfy the physical and psychological needs of family members. However, in its urban setting, the nuclear family is fragmented from its kinship network, leading to psychological isolation. Distancing itself from the extended family results in loss of its productive, political, and religious functions. Social mobility, particularly in the highly mobile U.S. culture, was made possible by the breaking of family ties.

However, most sociologists, anthropologists, and psychologists today disagree with Parsons' depiction of the nuclear family as isolated from its kin. Evidence from studies of family networks indicates that ties are maintained between members of the nuclear family and kin at different generational levels in the United States and many European countries. Many ethnic groups in the United States maintain fairly close family and kinship ties, whereas others do not. Cross-cultural studies have indicated that there are not polar opposites of no family ties versus close family ties but, rather, different degrees of family ties. In countries such as the United States, Sweden, and Britain there are looser ties between the nuclear family and kin, whereas the ties are closer in countries such as China, India, and Greece.

Family theories based on systemic concepts have also been developed (Parson's theory is also a systems theory). Family system theory has been developed within general systems theory, and its basic concepts apply to physics, biology, and social sciences, among other sciences. The family is a system within larger suprasystems of society and attempts to maintain equilibrium by adapting to demands and to changes in the larger system. The focus is on communication processes between family members and on recurring family transactional patterns. Family system theory was developed in the fields of clinical psychology and psychiatry as a psychotherapeutic method. This has resulted in the construction of models of family functioning and techniques of psychotherapy directed toward the family. Ecological theories of family trace the relationships between environment, social institutions, the family, and psychological variables. The human ecology theory of Bronfenbrenner is basically a family systems model. One of the earliest ecological models is that of cultural anthropologists John and Beatrice Whiting and Irvin Child and their many collaborators in their six-culture study. Their basic hypothesis was that child training throughout the world is in many important aspects identical and that there are universal

aspects of behavior but also differences from culture to culture. Six cultures were studied in detail: those of Raira, Okinawa; Tarong, the Philippines; Khalapur, India; Nyansongo, Kenya; Juztlahuraca, Mexico; and Orchard Town, United States. The elements of the model were the environment (climate, flora, fauna, and terrain), history of the settlements, maintenance systems (subsistence patterns, means of production, settlement patterns, social structure, etc.), the child's learning environment (caretakers and teachers, mothers' workload, etc.), psychological variables (behavioral styles, skills and abilities, values, conflicts, and defenses), innate needs, and projective expressive systems or elements of the culture (religion, magic beliefs, ritual and ceremony, etc.) (Fig. 1). Cultural features such as the economy, social structure, settlement pattern, and household and family organization were found to be related to behavior. For example, the complexity of the socioeconomic system and the composition of the household were predictive of social behavior of the system. This is a classic cultural anthropological model for cross-cultural studies of family.

Other family theories employed include feminist theory, symbolic interactionism, family development, phenomenology, family power, and exchange theory.

3. FAMILY

3.1. Definition of Family

Currently, there are many definitions of family based on different theoretical perspectives. An early definition of the family was that of the anthropologist Murdock: "The family is a social group characterized by common residence, economic cooperation, and reproduction. It includes adults of both sexes, at least two of whom maintain a socially approved sexual relationship, and one or more children, own or adopted, of the sexually cohabiting adults." The functions of family were considered to be sexual, reproduction, socialization, and economic. The sociologist Popenoe defined family in terms of recent changes in the social and economic position of the family, the increase in one-parent families [divorced, adoptive, widows(ers), or unmarried mothers], and homosexual families. In this definition of the family, the minimum number is one adult and one dependent person, the parents do not have to be of both sexes, and the couple does not have to be married. The functions of the family are procreation and socialization of children, sexual regulation, economic cooperation, and provision of care, affection, and companionship.

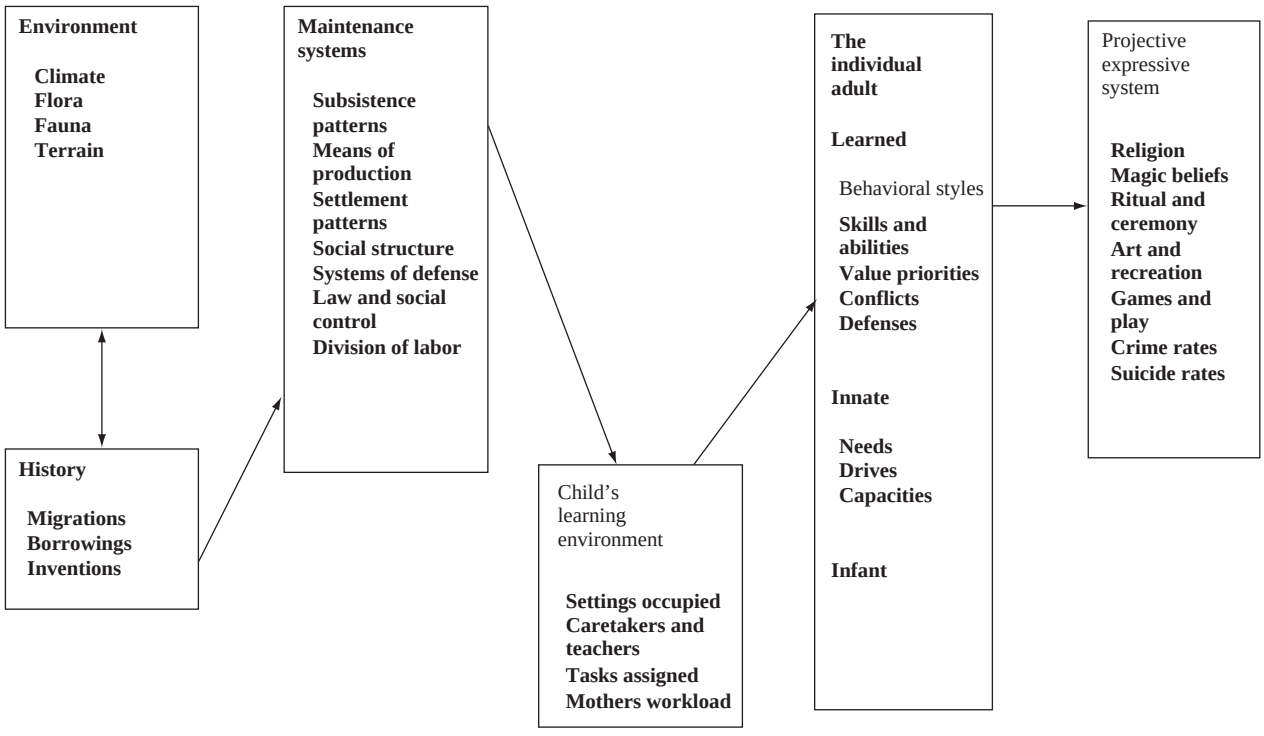


FIGURE 1 A model for psychocultural research. From Whiting and Whiting (1975).

3.2. Family Typology

As inferred in the previous definitions, there are different types of families. The structure refers to the positions of the members of the family (e.g., mother, father, daughter, grandmother, etc.) and the roles assigned to the family members by the culture. For example, traditional roles of the nuclear family in North America and northern Europe in the mid-20th century were the wage-earning father and the housewife and child-raising mother. Cultures have social constructs and norms related to the proper roles of family members—that is, what the role of the mother, father, etc. should be.

Family types or structures have been delineated primarily by cultural anthropological studies of small cultures throughout the world. However, family sociologists have also contributed to the literature on family typology, although sociology has been more interested in the European and American family and less interested in small societies throughout the world.

There are a number of typologies of family types, but a simple typology would be the nuclear and the extended family systems. To these can be added the one-parent family.

The nuclear family consists of two generations: the wife/mother, husband/father, and their children. The one-parent family is also a variant of the nuclear family. Most one-parent families are divorced-parent families; unmarried-parent families comprise a small percentage of one-parent families, although they have increased in North America and northern Europe. The majority of one-parent families are those with mothers.

The extended family consists of at least three generations: the grandparents on both sides, the wife/mother and the husband/father, and their children, together with parallel streams of the kin of the wife and husband. There are different types of extended families in cultures throughout the world. The following is one taxonomy:

- The polygynous family consists of one husband/father and two or more wives/mothers, together with their children and kin. Polygynous families are found in many cultures. For example, four wives are permitted according to Islam. However, the actual number of polygamous families in Islamic nations is very small (e.g., approximately 90% of fathers in Qatar, Kuwait, United Arab Emirates, Oman, Bahrain, and Saudi Arabia have only one wife). In Pakistan, a man seeking a second wife must obtain permission from an arbitration council, which requires a statement of consent from the first wife before granting permission.
- In a few societies in Central Asia there are polyandrous families, in which one woman is married to several brothers and thus land is not divided. However, this is a rare phenomenon in cultures throughout the world.
- The stem family consists of the grandparents and the eldest married son and heir and their children, who live together under the authority of the grandfather/household head. The eldest son inherits the family plot and the stem continues through the first son. The other sons and daughters leave the household upon marriage. The stem family was characteristic of central European countries, such as Austria and southern Germany. The lineal or patriarchal family consists of the grandparents and the married sons. This is perhaps the most common form of family and is also found in southern Europe and Japan.
- The joint family is a continuation of the lineal family after the death of the grandfather, in which the married sons share the inheritance and work together. Joint families were found south of the Loire in France, as were patriarchal families, whereas the nuclear family was predominant north of the Loire. Joint families are also found in India and Pakistan.
- The fully extended family, or the *zadruga* in the Balkans countries of Croatia, Bosnia, Serbia, Montenegro, Albania, Macedonia, Bulgaria, had a structure similar to that of the joint family but with the inclusion of cousins and other kin. The number of kin living and working together as a family numbered in the dozens.

A point needs to be made regarding the different types of extended families. Historical analyses of the family by anthropologists and sociologists indicated that people considered to be members of a family or a household were not necessarily kin. For example, in central European countries until the 18th century, servants (who were often relatives), semipermanent residents, visitors, workers, and boarders were considered to be members of the household. The term *familia* was used to denote large households rather than "family" in the modern sense. Until the 18th century, no word for nuclear family was employed in Germany but the term "with wife and children." Frédéric Le Play, considered to be the father of empirical family sociology, discussed the emergence of the nuclear family as a product of the industrial revolution. He also characterized the nuclear family, the *famille*, as unstable in comparison with the stem family.

One theory regarding the change from feudal familia to the famille of Western Europe is based on

the following analysis. After the reformation, vassals left the feudal towns to seek work in the cities. This led to the separation of the dwelling place and place of work and resulted in privacy and the sentimentality of the nuclear family. This pattern, however, was not found among the peasants in the agricultural areas. The strengthening of the relationship between parents and children was also a result of the religious influence of the Age of Enlightenment. These changes led to the releasing of servants from the close community of the household. Servants and workers became less personal and part of the household and more contractual. This led to the emergence of many new nuclear families (e.g., those of early factory workers and clerks). A new word in German, *Haus*, referred only to those living within it.

Historical analyses of the family during this period in Western Europe also emphasize that not all families were large extended families because establishing this type of household was dependent on land ownership. Most families worked for large feudal types of households and were essentially nuclear in structure. In England during this period, where land ownership was restricted to the nobility, the vast majority of families, which either worked for the landowners or rented small plots, were necessarily nuclear families.

3.2.1. The Nuclear Family: Separate or Part of the Extended Family?

The key element in studying different types of family structure and its relationships with psychological development of the children, its economic base, and its culture is the nuclear family. In 1949, Murdock made an important distinction regarding the relationship of the nuclear family to the extended family: "The nuclear family is a universal human social grouping. Either as the sole prevailing form of the family or as the basic unit from which more complex familial forms are compounded, it exists as a distinct and strongly functional group in every known society."

Murdock made an important point: The nuclear family is prevalent in all societies, not necessarily as an autonomous unit but because the extended family is essentially a constellation of nuclear families across at least three generations. Parsons' theory that the adaptation of the family unit to the industrial revolution required a nuclear family structure resulting in its isolation from its traditional extended family and kinship network, leading to psychological isolation and anomie, has had a strong influence on psychological and sociological theorizing about the nuclear family. However, studies of social networks in North America and northern Europe have shown that the hypothesized isolation of the nuclear family is a myth. Nuclear families, even in these industrial countries, have networks with grandparents, brothers and sisters, and other kin. The question is the degree of contact and communication with these kin, even in nations of northern and southern Europe.

A second issue relates to the different cycles of family, from the moment of marriage to the death of the parents or grandparents. The classic three-generation extended family has a lifetime of perhaps 20–30 years. The death of the grandparent, the patriarch of an extended family, results in one cycle closing and the beginning of a new cycle with two or three nuclear families, the married and unmarried sons and daughters. These are nuclear families in transition. Some will form new extended families, others may not have children, some will not marry, and others (e.g., the second son in the stem family) will not have the economic base to form a new stem family. That is, even in cultures with a dominant extended family system, there are always nuclear families.

A third issue is the determination of a nuclear family. This is related to place of common residence or the "household" of the nuclear family. Demographic studies of the family usually employ the term household in determining the number of people residing in the residence and their roles. However, there is a paradox between the concepts household and family as employed in demographic studies. Household refers to counting the number of persons in a house. If there are two generations, parents and the children, they are identified as a nuclear family. However, this may lead to erroneous conclusions about the percentage of nuclear families in a country. For example, in a European demographic study, Germany and Austria had lower percentages of nuclear families than Greece. This appears to be strange because Greece is known to be a country with a strong extended family system. However, demographic statistics provide only "surface" information, which is difficult to interpret without data about attitudes, values, and interactions between family members. Nuclear households in Greece, as in many other countries throughout the world, are very near to the grandparents—in the apartment next door, on the next floor, or in the neighborhood—and the visits and telephone calls between kin are very frequent. Thus, although nuclear in terms of common residence, the families are in fact extended in terms of their relationships and interactions.

In addition, there is the psychological component of those who one considers to be family. Social representation of his or her family may consist of a mosaic of parents, brothers and sisters, grandparents, uncles, and aunts and cousins on both sides, together with different degrees of emotional attachments to each one, different types of interactions, bonds, memories, etc. Each person has a genealogical tree consisting of a constellation of overlapping kinship groups—through the mother, father, mother-in-law, father-in-law, but also through the sister-in-law, brother-in-law, cousin-in-law, etc. The overlapping circles of nuclear families in this constellation of kin relationships are almost endless. Both the psychological dimension of family—one's social representation—and the culturally specified definition of which kin relationships are important determine which kin affiliations are important to the individual ("my favorite aunt") or the family ("our older brother's" family) and which are important in the clan (the "Zaman" extended family) or community (the "Johnsons" nuclear family). Thus, it is not so important "who lives in the box" but, rather, the types of affiliations and psychological ties with the constellation of different family members or kin in the person's conception of his or her family, whether it is an "independent" nuclear family in Germany or an "extended family" in Nigeria.

3.3. Subsistence Patterns and Family Types

Ecological theories specify relationships between the family structure and function, the ecological context (physical environment, climate, flora, fauna, terrain, and settlement patterns), the sociopolitical context (social structure, subsistence patterns, traditions, religion and beliefs, language and means of communication, etc.), and psychological variables (cognition, personality, values, social interactions, etc.). The ecological context and the sociopolitical context can shape family types, but the model is not strictly deterministic: Individuals or the community can also shape family types. The number of wives, the dowry system, and relationships with kin in a family type can be shaped by traditions and religion. However, perhaps the most important determinant of family type is the subsistence system, which is also shaped by the ecological context, particularly in small societies. In hunting and gathering societies, in which mobility of the family is necessary in order to obtain food, the nuclear family type is prevalent. In agricultural and pastoral societies, in which there is a permanent base and many hands are necessary for cultivation of the land, the extended family types are prevalent. In modern industrial society, in which people are hired to provide services or to work in industries, money is a means of exchange, people live in urban areas, and apartments or houses are expensive, the nuclear family is prevalent.

4. KINSHIP PATTERNS AND FAMILY MATTERS

4.1. Kinship Relationships

Cultural anthropology has categorized the different types of kinship relationships in extended families in a variety of cultures. These are called lineal, collateral, and affinal relationships. Lineal relationships are those between the grandparents and the grandchildren. Collateral relationships are those with uncles and aunts, cousins, and nephews and nieces. Affinal relationships are those between parents-in-law, children-in-law, and siblings-in-law as well as with matrilineal and patrilineal kin. However familiar the terms for these kin are in Europe and North America, they are critical concepts in different cultures because they are also related to the types of relationships and obligations toward affinal, collatoral, and affinal kin, to lines of descent, to residence, to inheritance of property, and to roles.

The terms for kin are more differentiated in various cultures. Murdock employed a system with six categories based on categorization of cousins: Hawaiian, Eskimo, Sudanese, Iroquois, Crow, and Omaha. Other categorizations are based on classification of paternal and maternal relatives. For example, in Iran the term in Farsi for mother's sisters is *khaleh* and it is *ammeh* for father's sisters, with analogous terms for father's brothers. In Pakistan, several cousins may be raised together in the same household. They are referred to in Urdu as brothers or sisters, with the prefix added, "through-my-maternal-aunt" or "through-my-paternal-uncle." Generic words, such as aunt, uncle, or even grandparents, have no equivalent in Urdu, but specific terms related to matrilineal kin, as in Iran, are employed; for example, *khala* is the mother's sister and *phuphee* the father's sister.

4.2. Residence

Residence refers to where the couple resides after marriage. One theory is that subsistence patterns

determine rules of residence, which in turn affect the rules of descent. The theory is based primarily on small, noncomplex societies.

- Patrilocal refers to residence with or near the husband's patrilineal kinsmen. This is the most common form of postmarital residence, with approximately 50% of small societies sampled.
- Virilocal refers to residence near the husband's father and is similar to patrilocal residence. The difference is the absence of patrilineal kin groups. This represents the second largest percentage (20%) of postmarital residence after patrilocal. The remainder of residence patterns sum to the remaining 30% of societies.
- Matrilocal is similar to patrilocal but refers to residence with or near the wife's matrilineal kinsmen. This is the case, for example, with the Chiapas of Mexico.
- Avunculocal refers to residence with or near the maternal uncle or other male matrilineal kinsmen of the husband.
- Bilocal is residence with or near the parents of the husband or the wife.
- Matrilocal–avunculocal is similar to bilocal but can be either matrilocal or avunculocal residence.
- Avunculocal-virilocal is also similar to bilocal but can be either avunculocal or virilocal residence.
- Neolocal means residence apart from the relatives of both spouses. This is characteristic of nuclear family residence in northern Europe and North America. In complex industrial societies in which there are few small independent farms, the economy is based on wages rather than land ownership. Thus, land ownership as a basis of production is not a factor, and small nuclear families with neolocal residence are more functional.
- Duolocal means separate residence of each spouse in the paternal and maternal household.

4.3. Descent

Descent refers to the relationships with paternal and maternal kin. The relationships concern many aspects of life, such as economic activities, religious activities, child-rearing practices, political behavior, inheritance, and residence. These aspects of living are organized around strict sets of rules with kin, many of them as ritualized behaviors. Societies with specific rules for different kin are usually hierarchical in structure and in the distribution of power. The following are rules of descent in different cultures.

Bilateral: Affiliation is with both mother's and father's relatives.

Patrilineal: Affiliation is with kin of both sexes through the maternal and paternal fathers only, but not through maternal and paternal mothers.

Matrilineal: Affiliation is with kin of both sexes through the maternal and paternal mothers only, but not through maternal and paternal fathers.

Ambilineal (cognatic): Affiliation is with kin through either the maternal parents or the paternal parents. Some may affiliate with kin through their mothers and others through their fathers.

Double: Affiliation is with both father's patrilineal kin and mother's matrilineal kin.

4.4. Marriage

Societies have norms regarding whom one is permitted to marry (endogamy) and restrictions regarding whom one cannot marry (exogamy). In some societies, such as in India or Pakistan, endogamy means that marriage is restricted to the same caste, the same village, the same religion, and the same race. These social norms are not as restrictive in North America and Europe, although marriage to someone of another racial or ethnic group or to someone of a different level of education or social status may not be accepted in some groups. Exogamy means that, e.g., marriage to a member of the nuclear family, or incest, is a universal taboo in almost all societies. In China, marriage to someone with the same family name was disapproved. In many societies, marriage is not permitted with first cousins or with the son or daughter of a godparent. In other societies, such as Saudi Arabia, marriage with first cousins is desired, specifically to the sons and daughters of the paternal brother (uncle) or to members of the clan. In other societies, if a husband dies, the wife must marry the brother.

In the past, in most societies throughout the world, marriages were arranged between two families, and a verbal or written contract was agreed upon regarding the dowry or the bride wealth; however, this is changing gradually in many societies in Africa and Asia. In patrilineal societies, the fathers usually made the decision. In the few matrilineal societies in the world, the mother's brothers made the decision. Decisions to unite their sons and daughters in marriage are based on certain criteria. Because endogamy is an important restriction, families look for the best possible alliances between them, which may be based on social and economic factors and, in many cases (such as in Saudi

Arabia, where marriage between paternal first cousins is highly desirable), strengthening the wealth, prestige, and social power of the extended family within the clan or tribe. The decision of a young couple to marry on the basis of love with or without the consent of the parents occurs in a minority of societies, including North America and northern Europe, although this is changing rapidly in many societies. One theory is that in societies with neolocal postmarital residence, romantic love as a basis of marriage is more common. It is suggested that marriage based on love and companionship provides a stronger union when the couple must depend on their own economic assets and not on their parents' aid for survival.

The dowry is an integral part of the marriage contract between the two families of the couple. There are three general types of dowries. The dowry system refers to the transfer of property, money, or animals by the bride's parents to the husband after marriage. The bride wealth system, the most common dowry system, is the opposite of the dowry system: The transfer is from the husband's family to the bride's family. A third type is bride theft, in which the man abducts the woman without the consent of the respective families. In the past, the dowry system was in effect in European countries, in addition to other societies, whereas the bride wealth system was in effect in Islamic nations as well as other societies.

Inheritance of property is an integral part of marriage and lineal descent. Japan and China adapted differently to industrialization in the 19th century. Both had an agrarian economy, rapid population growth, and extensive but corrupt and inefficient bureaucracy. In China, inheritance was egalitarian, but in Japan a single child inherited the property, which made it possible to accumulate capital. In China, loyalty was to the family and nepotism a duty, so every family member could benefit from upward social mobility. In Japan, which had a more feudal system, a father could disinherit his son and adopt a young man who seemed more worthy.

4.5. Divorce

There is considerable cultural diversity regarding divorce. Divorce is socially disapproved in all societies, and families usually make great efforts to keep the couple together. The reasons for this are many: the social stigma of divorce, the consequences for the children, the financial provisions, and the issues of property and inheritance.

Societies differ in the degree to which divorce is controlled and by which institutions. In some societies, divorce is controlled directly by the family, whereas in others there is indirect control by social institutions and by the dominant religion. For example, Catholicism does not permit divorce except under highly unusual situations requiring a special dispensation. The Orthodox Church permits three marriages and three divorces. Islamic law, the *sharia*, permits polygamy up to four wives and also divorces. However, in Islamic nations, the intent of a husband to take a second wife may lead to the first wife seeking a divorce. Despite the conventional notion in the West that in Islam husbands can divorce a wife by stating three times the phrase, "I divorce you," divorce is not an easy matter because of its legal and social consequences. According to Islamic law, the daughter inherits property from the father—half of the sum of the son's. The wife retains property in her name after marriage, even gifts from her family, and the husband has no legal claim to her property. After divorce, the woman retains her property and custody of the children until they reach 7 years of age. However, the father remains the legal guardian and is responsible for the financial care of the children. Remarriage of a divorced woman is permitted, but the children of the first marriage are subsequently raised by the maternal grandmother.

5. FAMILY ROLES AND POWER

Role theory refers to the cultural norms regarding psychological and interactional aspects of members of society, such as mothers, fathers, sons, daughters, and grandparents. The originators of role theory are Ralph Linton in sociology and George Herbert Mead in social psychology. Role refers to the social expectations and the social scripts of family roles—how roles have been shaped by cultural conventions and by the collective ideologies of a society. One aspect of role theory studies how roles are learned during the process of social interaction. That is, people interact with others, they see themselves and others as occupants of particular statuses, and they learn guides for action. In other words, there are certain social scripts or expectations associated with certain roles.

Structural role theory as developed by Linton and Parsons refers to the structural and functional aspects of status and role. By structure is meant the positions recognized by the specific social system and the content of the role. That is, what are the social scripts associated with the roles of mother and father in the society? For example, social scripts of gender roles in a society might

be that the place of mother is in the home, she should raise the children, cook, etc., whereas in other societies mothers should work, share with the husbands the daily work of the home and care of children, etc. The functional analysis is concerned with the function of role. That is, what are the consequences for the rest of the social system of the specific role? How does the role contribute to the maintenance of the system? How does a particular role help in the achievement of the goals of the system? How does the role help solve the process of adaptation to the social system and contribute to the maintenance of patterns, such as the values system? Structural role theory also has a comparative element, attempting to determine which structure features of social systems are universal through cross-cultural analysis.

Power, the degree of gender equality, and the division of resources within the family are another aspect of family dynamics in the family. Anthropology has established that power is related to the control of the economic activity or land ownership and institutionalized through the norms of the society. Power can be defined as the influence of the spouse, children, and other relatives. However, the influence is not a personality trait, although there may be individual differences in the ability to influence others. The basis of the influence is the authority bestowed by social norms to different family roles and internalized in terms of values. Power may be expressed in the control of the finances, in the ability to command respect and obedience, in feelings of autonomy and control over one's life, in the capacity to regulate one's dependency on others, in the power to decide where the family lives and whose career is given priority, and in the freedom to be able to leave difficult family situations.

Most agricultural societies are patrilineal, but some are matrilineal. In some agricultural societies, the male does most of the plowing or animal tending and the woman's role is to care for the house and the children. However, in many patrilineal agricultural societies, the woman does much of the work, such as harvesting and tending the crops, in addition to the household tasks and caring for the children. Boys learn the roles of the males and the fathers, and girls learn the roles of the females and the mothers in small societies. The roles in small societies are hierarchical, in which the patriarch of the family has the ultimate power and males have more power than females. In matrilineal societies, the matriarch has more power than the father. The brothers of the matriarch have more power than the father, who is absent from the family and may visit the matriarch secretly at night. However, although the father may have the institutionalized social power within the family, the influence of the mother in extended families regarding many family matters may be greater than that of the father.

It is important to note that role theory refers to social values and social scripts—what the social scripts should be in the society. However, there is also a psychological dimension; there are individual differences in each society regarding agreement with these roles. Some members of society insist on the strict application of these roles, whereas some members rebel against these roles. In many societies throughout the world undergoing transition, women are rejecting the traditional roles of mother, housewife, and caretaker and are entering the workforce, as in Western societies. However, the process is more painstaking in these societies, and working mothers often experience strong family pressures, which may come from the grandparents or the in-laws, to be both a working mother and fulfill the traditional roles of housewife and caretaker. The structure of the traditional family in many countries provides an alternative caretaker of the child while the mother works—the grandmother.

6. PSYCHOLOGICAL DIMENSIONS

So far, the discussion has covered features of culture and family. Societies differ in terms of degree of complexity, from small, simple societies of a few hundred members who subsist through fishing or hunting to large, complex societies of western Europe and North America in which the economic basis is industry and services. Societies also differ in terms of family structure and function. Family type is related to the subsistence system, property and residence, and descent. This section discusses whether there are different psychological characteristics related to different family types.

As indicated previously, there are relationships between means of subsistence and family types. Agricultural and pastoral societies tend to have extended families, whereas hunting, fishing, and gathering societies tend to have nuclear societies. Agricultural societies tend to have a base, land and permanent houses, and large families and kin nearby. A large number of kin is necessary for crop cultivation. Cultivation is a process that requires patience and the cooperation of many, and it is affected by droughts, floods, and other natural forces. Studies have found that children in agricultural and pastoral societies are taught to be responsible, compliant, obedient, and to respect their elders. Hunting and

gathering societies move about in order to obtain food, hunting requires particular skills, and good hunters are respected no matter what their age. Children in these societies tend to be self-reliant, independent, and achievement oriented. It is likely that these personality traits and the social values that promote them are a function of the adaptation of the individual and the family to the ecological constraints of their environment through their means of subsistence.

In the six-cultures studies of Whiting *et al*., relationships between the economy, social structure, settlement pattern, household and family organization, and the behavior of children were studied. Minturn and Lambert studied the effects of differences in living patterns and economic activities on child-rearing practices and the behavior of children (e.g., succorance, achievement, self-reliance, obedience, nurturance, responsibility, sociability, dominance, and aggression). Mothers were least warm in Mexico and India, where the living patterns were characterized by courtyard or semicourtyard living. In African and Philippine societies, where mothers lived in separate houses surrounding a partially enclosed yard, mothers were intermediate in warmth. Okinawan mothers, who had more privacy, were the warmest. New England mothers, who had the most privacy and almost exclusive responsibility for child care, ranked lowest in warmth. In general, children in nuclear, joint, and lineal family households were treated more warmly and were indulged when alternate caretakers, such as grandparents, sisters, and aunts, were available.

Kagitcibasi's model of family change based on socioeconomic development posits three family interaction patterns: the traditional family in developing countries characterized by total interdependence between generations in material and emotional realms, the individualistic nuclear family model of Western society based on independence, and a dialectical synthesis of these two involving material independence but psychological interdependence between generations. The parenting orientations differ significantly among these patterns, as do the distinctive characteristics of the emergent selves. With increasingly urban lifestyles and increasing affluence in collectivistic cultures, material interdependence between generations decreases because elderly parents no longer depend on the economic support of their adult offspring, but emotional interdependence continues. This leads to increased autonomy in child rearing, but together with psychological interdependence.

These changes are theorized to lead to the development of three types of the self in the three family models: the related self, high in relatedness but low in autonomy in the family model of total interdependence; the separated self, high in autonomy but low in relatedness in the family model of independence; and the autonomous–related self, high in both relatedness and autonomy in the family model of psychological interdependence, revealed in the changing family in collectivistic societies.

The previously discussed studies refer primarily to personality and emotional variables. Family type is also related to cognitive variables. An example is Berry's ecocultural framework, which traces the relationships between ecological, cultural, and acculturational variables and cognitive styles. Initially, links were established between hunting-based and agriculture-based peoples regarding ecological demands and cultural practices and the required cognitive performances (i.e., visual disembedding and analytic and spatial abilities). These links were analyzed in detail in terms of the variables subsistence patterns, settlement patterns, population density, family type, social/political stratification, socialization practices, education, and wage employment. It was found that nomadic hunters and gatherers, with a loose social structure, nuclear family systems, and an emphasis on assertion in socialization, were relatively field independent (i.e., were able to focus on an object without being influenced by its context), whereas sedentary agriculturalists, with tight social structures, extended family systems, and an emphasis on compliance in socialization, were relatively field dependent (i.e., strongly influenced by the context of stimuli). In addition, people in societies undergoing acculturation and developing Western types of education became more field independent (Fig. 2).

7. CONCLUSION

This article presented a picture of the family primarily from the perspective of cultural anthropology, sociology, and cross-cultural psychology. In this brief synopsis, it was not possible to present the contributions of other social sciences, such as economics, geography, psychiatry, and political science. The primary goal was to focus on the study of family from the standpoint of cross-cultural psychology and psychological anthropology. Thus, emphasis was placed on an ecological or ecocultural perspective: the ecological context, social structure, maintenance systems, means of subsistence, the family, child development, roles, and psychological consequences. The picture may appear to be like a kaleidoscope in that there is wide cultural variety and there are many variations of families and social structure.

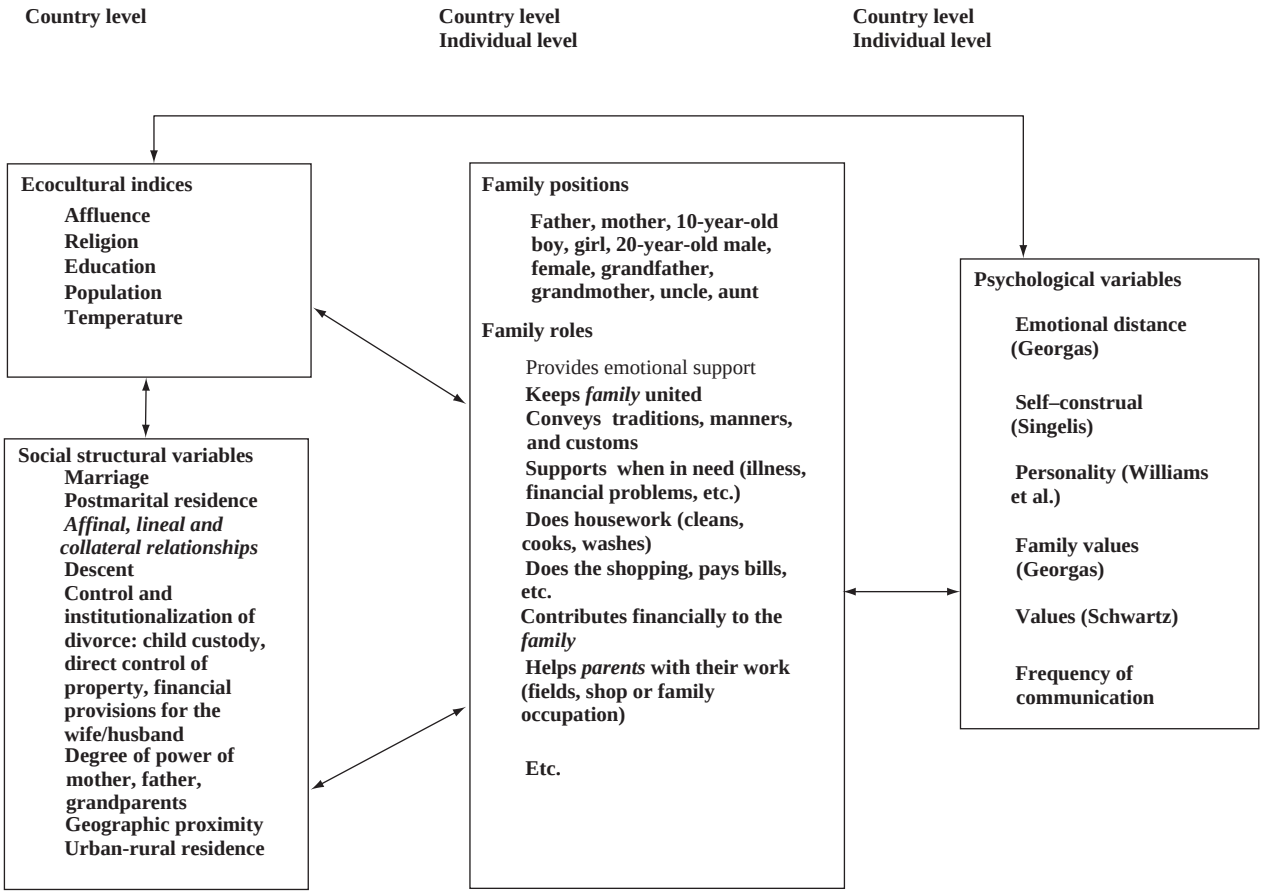


FIGURE 2 An ecocultural framework for the study of family.

However, there have been many efforts to determine the structure underneath this seemingly chaotic plethora of findings. This entails a comparative approach, consistent with the approach of psychological anthropology and cross-cultural psychology in seeking universals in family structure and function as well as variations due to cultural differences. A criticism of sociology, which has made significant contributions to the study of the family, is that it has been concerned primarily with the study of the family, particularly the nuclear family, only in the countries of northern Europe and North America.

An important issue in making sense of results from thousands of societies is the chronological dimension. That is, anthropologists have been studying small societies for more than a century. Thus, some findings refer to the study of a culture 100, 80, 40, or 30 years ago and some to recent studies. However, in a changing world in which small societies have been exposed to economic changes, technology, television, tourism, and trade with economically developed nations of the West and Asia, these societies have also been adapting. Acculturation and enculturation in response to these pressures for change have also affected the links between ecology, social structure, family types, and psychological variables. There is a trend toward more families becoming structurally nuclear, even in small societies. However, one must be careful in employing the phrase "becoming nuclear." There is increasing evidence from studies of small societies or developing nations that the predictions of modernization theory that the driving force of modernization, the economic engine, will eventually lead to the Western type of nuclear family system may be incorrect. The number of nuclear families is increasing in many developing societies; the traditional family systems are no longer totally dependent on subsistence systems, such as hunting, gathering, or even agriculture; young couples are increasingly choosing their spouses rather than entering arranged marriages; women are increasingly entering the workforce; traditional family roles have changed; and the male is no longer the

absolute patriarch in the family. However, even though the number of nuclear families is increasing in most societies, these families still maintain very close relationships with their kin. It appears that although the structure of the family is changing from extended to nuclear, the functions of the family and the kin relations have not changed as much. Interactions and relations with kin are the psychological components of the family, which are of particular concern to psychologists.

The study of culture, family, and the psychological consequences is thus in a very dynamic phase and of considerable interest to social scientists and students.

See Also the Following Articles

Marital Therapy

Further Reading

Berry, J. W., Poortinga, Y. H., Segall, M. H., & Dasen, P. R. (2002). *Cross-cultural psychology: Research and applications*. New York: Cambridge University Press.

Bronfenbrenner, U. (1979). *The ecology of human development*. Cambridge, MA: Harvard University Press.

Georgas, J., Berry, J. W., van de Vijver, F. J. R., Kagitcibasi, C., & Poortinga, Y. H. (Eds.). (2005). *Cultures, family and psychological functioning*. Cambridge, UK: Cambridge University Press.

Ingoldsby, B. B., & Smith, S. (Eds.). (1995). *Families in multicultural perspective*. New York: Guilford.

Levinson, D., & Malone, M. J. (1980). *Toward explaining human culture*. New Haven, CT: HRAF Press.

Levi-Strauss, C. (1969). *The elementary structures of kinship*. London: Eyre & Spottiswoode.

Malinowski, B. (1962). *Sex, culture, and myth*. New York: Harcourt, Brace, & World.

Minturn, L., & Lambert, W. W. (1964). *Mothers of six cultures*. New York: Wiley.

Munroe, R. H., Munroe, R. L., & Whiting, B. B. (1981). *Handbook of cross-cultural development*. New York: Garland.

Murdock, P. M. (1949). *Social structure*. New York: Free Press.

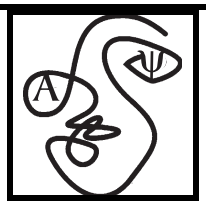
Parsons, T. (1943). The kinship system of the contemporary United States. *American Anthropologist*, *45*, 22–38.

Parsons, T. (1949). The social structure of the family. In R. N. Anshen (Ed.), *The family: Its functions and destiny*. New York: Harper.

Popenoe, D. (1988). *Disturbing the nest—Family changes and decline in modern societies*. New York: Aldine.

Segalen, M. (1986). *Historical anthropology of the family*. Cambridge, UK: Cambridge University Press.

Whiting, B., & Whiting, J. W. M. (1975). *Children of six cultures: A psycho-cultural analysis*. Cambridge, MA: Harvard University Press.



Feminist Psychology

María Inés Winkler Müller
Universidad de Santiago (USACH), Santiago, Chile

1. Historical Background
2. Psychoanalysis and Feminist Psychology
3. Feminist Psychology and Moral Reasoning
4. A Feminist Ethic
5. The Feminist Perspective in the History of Psychology
6. Clinical Psychology and the Feminist Perspective
7. Community Psychology and the Feminist Perspective
8. Feminist Psychological Methodology
9. Conclusion
Further Reading

GLOSSARY

feminist ethics A philosophical stance carrying the assumption that women and their experience have moral meaning, focusing attention on feminine ethical values such as the ethic of care, implying a critique of all discriminatory distortions, and promoting the necessary social action to get rid of all kinds of oppression.

moral reasoning The ways in which people analyze and resolve ethical questions.

psychoanalytic critical posture Theoretical stance based on women's libidinal and genital differences from men.

Feminism is mainly a political, social, and cultural perspective that is pro-women. That is, it is an ideology or an intellectual posture that attempts to explain why things are the way they are while advocating for political, social, and cultural equality between women and men. It also tries to eradicate gender discrimination in society.

Feminist psychology refers both to a collection of efforts to understand the different ways in which women and women's issues have been considered (or not considered) in psychology and to a body of knowledge about women gathered since the 1960s. Feminist psychology, also known also as psychology of women, analyzes the process through which gender is constructed, emphasizing the worth in women's perspectives and experiences. One can recognize different perspectives in feminist psychology, but all of them poignantly signal and challenge the invisibility of women in psychology. Furthermore, it is a double exclusion; women are not considered, and one does not even recognize that there is an omission. Feminist psychology challenges the validity of gender role stereotypes and expectations that have traditionally maintained patterns of male dominance and female subordination in our society. More specifically, feminist psychology disputes unquestioned premises, some research traditions, and theoretical positions that maintain a patriarchal order in which women remain secondary. At the same time, it questions those professional practices and institutional and societal structures that limit the understanding, treatment, professional achievement, and responsible self-determination of women and men.

1. HISTORICAL BACKGROUND

Historically, feminist psychology has emerged as a grassroots movement, but it is possible to trace its origins and

©2004 Elsevier Inc.
All rights reserved.

identify some landmarks in the United States, especially in the works of Weisstein and Gilligan, where they denounced prevailing sexist biases and women's nonexistence in psychology. These pioneers emphasized that there were no female authors in any of the major theories of human development and psychological functioning and that women have been missing and ignored, even as research participants, at the formative stages of psychological theories. As Weisstein pointed out in 1971, "psychology is men's fantasy." That is, scientists draw their conclusions from investigating men and blindly generalize men's experiences to the experiences and lives of women but not vice versa. Weisstein's article, "Psychology Constructs the Female," challenged the predominant sexism in psychology while pointing out a double criticism of mainstream psychology: that it lacks scientific rigor and that it ignores the role played by social contexts. The lack of scientific rigor refers to the omission of the experiences of half the population (i.e., of women) in research that supports psychological theorizing. This lack of awareness about social context refers to the individualist emphasis that has characterized psychology in methodology as well as in theory construction, for example, ignoring the importance of power and gender in behavior. These criticisms are still valid, and many feminist psychologists reiterate this assertion. Today, the goal is still to do "improved science," taking into account all relevant variables and not just men's experiences. Cole's 1979 work was a pioneering study, titled *Fair Science,* that focused on sex-based discrimination and the stratification of science. Its main contribution was an increased understanding of women's challenges, opportunities, and obstacles when attempting a career in science.

Some interest and emphasis in the matters of sex differences started to appear from the beginning of psychology in the United States at the end of the 19th century and the beginning of the 20th century. Functionalism then grew as central theory strongly associated with prevailing Darwinian postures, especially with the variability hypothesis (men would have more variations among them than would women) and the functional variability hypothesis (women's intellectual performance would be affected by the menstrual cycle). These beliefs about a supposed female inferiority (despite the early disagreement and opposition found in the research done by Bradford Thompson, Puffer, Whiton Calkins, and other psychology female pioneers) prevailed for many decades as a subtext in psychological theorizing. This supposed inferiority showed itself in different social and professional fields, and a succession of studies developed during the 1970s and 1980s revealed that sexism was clearly present at several levels. Discrimination against women hindered women's access to, and success in, certain professions. Women encountered discriminatory practices when applying for graduate training, in hiring, and in publishing their research. In addition, research methodology and practice were characterized by male bias that influenced the selection of research design, assessment measures, and sex of participants studied as well as the analysis and interpretation of data.

Based on this new awareness, and because they were interested in achieving real changes in the theory, research, and practice of psychology, Weisstein, Chesler, Evans Gardner, and others in 1970 founded American Women in Psychology, now Division 35 of the American Psychological Association, created in 1973. In Canada, some of the most important events that marked the beginning of feminist psychology were the Underground Symposium, the Canadian Psychological Association (CPA) Task Force on the Status of Women in Canadian Psychology, the founding of the CPA Interest Group on Women and Psychology, and the establishment of the CPA Status Women Committee. These and other initiatives, mainly in the academic domain, contributed to highlight the presence and experiences of women in various subfields of psychology (e.g., moral reasoning, feministic ethics, history of psychology, clinical psychology, community psychology).

2. PSYCHOANALYSIS AND FEMINIST PSYCHOLOGY

Outside mainstream psychology, the feminist perspectives most known are related to psychoanalytic perspectives. Using her own mothering experience, Deutsch, one of the first woman analysts in the Vienna Psychoanalytic Society, underlined the consequences of the Freudian undervaluing of women in women's developmental psychological well-being in her pioneering text, *The Psychology of Women*. Horney, a psychoanalytic feminist, criticized Deutsch for equating women with masochism, challenging the role of gender in personality development while stressing the importance of society and culture as determining factors for normal and pathological development. Contemporary stances have also embraced a critical position, particularly object relations theory as developed by Chodorow and Benjamin as well as by French psychoanalysts such as Kristeva, Cixous,

and Irigaray. The French feminist project of "writing the body" is an especially revolutionary effort to provide women's writings with an authority based in women's libidinal and genital differences from men. Theories about the body are particularly important for feminists because historically the body has been considered equivalent to the feminine, the female, or the woman—and, as such, has been denigrated as weak, immoral, or unclean.

3. FEMINIST PSYCHOLOGY AND MORAL REASONING

In regard to moral reasoning, Kohlberg, in his well-known developmental research, studied the ways in which people resolve ethical questions and proposed the postconventional stage as the highest level of moral reasoning. This level is characterized by the adult's ability to base morality on the logic of principled decision making, which in turn is based on principles thought to be universal and not dependent on culture—the ethic of justice, in Kohlberg's words. But given that all of his research participants were male, Kohlberg could not have taken into account the different socialization of little girls and little boys in our culture. Males are traditionally socialized to be autonomous and independent, whereas females are supposed to be passive but loving caretakers of others. Gilligan argued that these differences in socialization practices lead to different values. For men, these differences lead to a morality based on equal rights and acceptance of abstract principles, even at the sacrifice of people's well-being. For women, these differences lead to a morality based on caring in which increasing maturity broadened the scope of their sense of responsibility and compassion. So, women have a different way—not an inferior one—of solving ethical dilemmas—the ethic of care. Olsen, in her 1978 book *Silences,* documented the history of women writers and pointed out the relative silence of women as literary voices. Gilligan, in her 1982 book *A Different Voice,* delineated one of the most well-known psychological theories about women's development. By listening to girls and women resolve serious moral dilemmas in their lives, Gilligan formulated the development of a moral reasoning organized around concepts of responsibility and care. She argued that the experiences of inequality and subordination that outline the lives of women and girls creates a moral sense of self that is grounded in human attachments and characterized by a strong connection with relationships. Gilligan and other representatives of the Harvard Group of Women Studies expressed it using the metaphor of "voice" and "silence." There were two voices, or two ways of experiencing oneself in relation to others, but only one voice—the male's—was acknowledged and delineated in the field of psychology and education. To give voice to women and listen to the unheard voice was the new motto. Gilligan recognized a deep affinity with the work of Baker Miller, and she was inspired by the radical consciousness regarding the situation of women as key for understanding the psychological order and voice central to women's experience and women's exercise of power. These American feminists saw women's strength emerging from the experience of recovering an authentic voice and the capacity to express it. Conversely, women's subordination was rooted in the silencing of their voice. But Gilligan's work has not been free of controversies. The question of sex differences is not always considered a feminist question, and her proposal of an alternative moral system reflecting women's experience can also be understood as a consequence of a male-biased culture.

4. A FEMINIST ETHIC

New developments in feminist psychology and in the domain of ethics suggest a feministic ethic that draws from women's experience and feminist theory to describe its principles. Feminist ethics embraces much of traditional ethics but goes beyond and provides unique lenses for determining what one ought to do and how to decide on courses of action in situations involving competing ethical principles. Because moral philosophy has been largely a male enterprise, feminists seek to expose the individual and institutional practices that have negated and devalued feminine thinking. Feminist ethics assumes that women and their experience have moral meaning, and it focuses attention on feminine ethical values such as the ethic of care. This perspective also implies that a feminist critique must be accompanied by a critique of all discriminatory distortions and by the necessary social action to get rid of all kinds of oppression and not just just women's oppression.

5. THE FEMINIST PERSPECTIVE IN THE HISTORY OF PSYCHOLOGY

In the history of psychology, mainly U.S. and Canadian psychologists denounced the androcentric bias in the

chronicling of the discipline. As with history, sociology, literature, and other disciplines, evidence shows that in psychology women have been systematically excluded as protagonists when historicizing the discipline. The growing awareness of this absence gave rise to a body of research that enriched the field and introduced the names of several female pioneers in the historic chronicle, as shown today in registers of the American Psychological Association and Canadian Psychological Association. In the United States, Bohan, Scarborough, Felipe Russo, and Furumoto have assumed the task of repositioning women in the history of psychology. In Canada, Pike and Wright, among others, have assumed this task. Bosch and Ferrer, from Spain, have developed multidisciplinary research that undertakes to reflect about psychological thought and to identify androcentrism and misogyny in the theories of Freud, Jung, and others. They also have formulated the idea of a nonandrocentric science. The reconstruction of female participation and contributions in various countries of Latin America is an unfinished task.

6. CLINICAL PSYCHOLOGY AND THE FEMINIST PERSPECTIVE

In clinical psychology, feminism has a strong hold in psychotherapy. Feminist psychotherapy emerges as a response to male dominance of mental health professions and institutions, and it objects to the role performed by the mental health establishment in maintaining social and power inequities between men and women as well as between members of the dominant culture and those of ethnic minority groups. Although feminist therapists use diverse therapeutic techniques, they have a common goal of raising awareness of the harmful effects of discrimination on clients, for example, the effects of job discrimination based on gender, race, or sexual orientation. Probably more important still is that feminist therapists dispute the tendency of traditional (male-centered) psychotherapies to catalog as normal those attributes identified with the male-dominated mainstream culture and to label as abnormal those characteristics associated with femininity and with minority cultures. Feminist therapists advocate that women's expressions of anger and distrust toward dominant group members can be seen as successful adaptive responses to prejudice and discrimination. In denouncing traditional psychotherapies for failing to recognize the harmful effects of sexism, racism, and heterosexism, feminist therapists assert that traditional therapies categorize any behaviors that do not fit the interests of the dominant group as problematic. In so doing, traditional therapies silence women. Feminist psychology also mandates that therapists explore, recognize, and understand their own social biases; that is, they should cultivate awareness of the influence and importance of those biases in their work as therapists. Feminist theorists assume that social inequality, rather than individual psychopathology, often plays a main role in creating and maintaining many of the difficulties presented by clients in psychotherapy, especially when those clients belong to the oppressed groups. Furthermore, the main process in feminist therapy is the search for clients' sources of power rather than the search for pathologies in people. Supervision, advocacy for victims of sexual harassment, and treatment of eating disorders also constitute important topics in feminist psychotherapy where gender is of utmost importance.

7. COMMUNITY PSYCHOLOGY AND THE FEMINIST PERSPECTIVE

Between community psychology and feminist psychology, we find notable convergence, not only in similar values but also in a common historical background, with both having emerged during the 1960s as part of broad movements for social justice and equality. Both perspectives share in criticizing victim-blaming ideologies and reject simplistic or restrictive standards of health, normality, and behavior. Respect for differences, distributive justice, and equality are some of the distinctive values that community psychology and feminist psychology endorse. Both perspectives dispute individual and adjustment-oriented analyses, explanations, and answers. Instead, both elaborate models and strategies emphasizing social policy, advocacy, prevention, and empowerment. In particular, the social community psychology represented by Latin American psychologists Montero and Serrano-García shares the commitment to the empowerment of the powerless, including social action and promotion of social change. Nevertheless, community psychology has not fully embraced feminist perspectives. It has been criticized for not adequately integrating women into the profession, not addressing women's issues in graduate training and

professional practice, and not incorporating feminist perspectives and processes into research, theory, and practice.

8. FEMINIST PSYCHOLOGICAL METHODOLOGY

In the methodological arena, feminist psychology states that more socially situated methods are needed because individualistic research methods have dominated feminist psychology and feminist research despite the theoretical command to study and understand people in context. It is suggested that the inclusion of more socially oriented methods, such as group interviews and focus groups, may help to overcome this deficiency. In addition, from a feminist ethics perspective, new questions arise about the role of power inequalities between the researcher and the researched and about the risk of the imposition of meaning. The task is to generate high-quality interactive data offering the possibility of theoretical advances through of the co-construction of meaning in which both the researcher and the researched participate.

Not only gender inspires different types of discrimination. In 1998, Burman denounced feminist psychology's exclusion of the experiences and voices of women whose cultural identities differ from those of White Western women. She included studies and reflections from contributors who lived and worked in diverse countries such as South Africa, Yugoslavia, and Puerto Rico. In these other cultural contexts, feminist psychology itself has only incipient development and has yet to influence mainstream psychology in any meaningful way.

From its beginnings, feminist psychology has continually expanded its scope of inquiry and its modes of studying gender. Over time, the frontiers that demarcate different approaches have become more visible. Postmodern feminist psychologists reject the possibility of a whole identity that includes a single true voice, and they sustain that diverse contradictory identities can have a libratory potential. On the other hand, in 2001, Marecek delineated what she termed "feminist skepticism," an approach that draws on postmodern thought and critical theory. The principal question is whose interest and ideas are served by psychology as science, assuming that psychology is not engaged in a harmless and disinterested search for the truth. Contemporary feminist theory calls into question the possibility of finding fixed, reliable, objective meanings that are applicable to everyone. New categories within current feminist analysis emerge, and differences among women are now in the foreground.

9. CONCLUSION

In synthesis, feminism has forced psychology to take a stand that includes women in theory as well as in research and practice. The introduction of women's psychology and girls' psychology as legitimate topics of study is no less important than the naming and exploring of main issues in women's lives and experiences. By now, feminist psychology has shown the power of women's voice in broadening our understanding of human development, and it is slowly succeeding in putting gender issues on mainstream psychology's agenda. Other dimensions of feminist psychology's contributions involve reconstructing research methods and priorities to study women in the context of their experiences, integrating multiple diversities into all subfields of the discipline, developing innovative approaches in psychotherapy, and transforming institutions toward being more inclusive and collaborative. Last, but not least, this perspective dictates advocacy for social action and public policies that benefit the health and well-being of both women and men.

Feminist's insights are transforming our understanding and the study of psychology and human development. Women's psychology faces future challenges—the call to address diversity among women, to promote and achieve developments in different countries and places, and to become more interdisciplinary as the discipline becomes more inclusive, especially in the task of integrating new ethnic and gender identities.

See Also the Following Articles

Gender and Culture ■ Gender and Education ■ Gender Role Development ■ Occupation and Gender

Further Reading

Baker Miller, J. (1987). *Toward a new psychology of women*. Boston: Beacon.

Belenky, M., Clinchy, B., Goldberger, N., & Tarule, J. (1997). *Women's ways of knowing: The development of self, voice, and mind*. New York: Basic Books.

Bohan, J. (1995). *Re-placing women in psychology*. Dubuque, IA: Kendall/Hunt.

Bond, M., & Mulvey, A. (2000). A history of women and feminist perspectives in community psychology. *American Journal of Community Psychology*, *28*, 599–630.

Bosch, E., Ferrer, V., & Gili, M. (1999). *Historia de la Misoginia* (History of misogyny). Barcelona, Spain: Anthropos.

Bosch, E., Ferrer, V., & Riera, T. (2000). *Una Ciencia no Androcentrica: Reflexions Multidisciplinars* (To a non-androcentric science: Multidisciplinary reflections). Palma, Spain: Universitat de les Illes Balears.

Brabeck, M., & Ting, K. (2000). Feminist ethics: Lenses for examining ethical psychological practice. In M. Brabeck (Ed.), *Practicing feminist ethics in psychology* (pp. 17–35). Washington, DC: American Psychological Association.

Brown, L. (2000). Feminist therapy. In C. R. Synder, & R. E. Ingram (Eds.), *Handbook of psychological change: Psychotherapy processes and practices for the 21st century* (pp. 358–380). Toronto: Wiley.

Burman, E. (1998). *Deconstructing feminist psychology*. London: Sage.

Chrisler, C. J., Golden, C., & Rozee, D. P. (1995). *Lectures on the psychology of women*. New York: McGraw–Hill.

Cole, J. (1979). *Fair science: Women in the scientific community*. New York: Free Press.

Deutsch, H. (1944). *The psychology of women. A psychoanalytic interpretation*. New York: Grune & Stratton.

Gannon, L., Luchetta, T., Rohes, K., Pardie, L., & Segrist, D. (1992). Sex bias in psychological research: Progress or complacency? *American Psychologist*, *47*, 389–396.

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development*. Cambridge, MA: Harvard University Press.

Henwood, K., Griffin, C., & Phoenix, A. (1998). *Standpoints and differences: Essays in the practice of feminist psychology*. London: Sage.

Horney, K. (1967). *Feminine psychology*. New York: Norton.

Kristeva, J. (1986). *The Kristeva reader*. New York: Columbia University Press.

Mahoney, M. (1996). The problem of silence in feminist psychology. *Feminist Studies*, *22*, 603–625.

Marecek, J. (2001). After the facts: Psychology and the study of gender. *Canadian Psychology*, *42*, 254–267.

Matlin, W. M. (1987). *The psychology of women*. New York: Holt, Rinehart & Winston.

Olsen, T. (1978). *Silences*. New York: Delacourte Press.

Pyke, S. (2001). Feminist psychology in Canada: Early days. *Canadian Psychology*, *42*, 4.

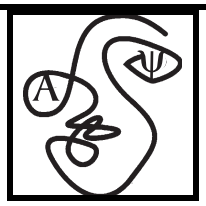
Unger, R. (1998). *Resisting gender: Twenty-five years of feminist psychology*. Thousand Oaks, CA: Sage.

Unger, R., & Crawford, M. (2000). *Women and gender: A feminist psychology*. New York: McGraw–Hill.

Weisstein, N. (1971). Psychology constructs the female. In V. Gornick, & B. K. Moran (Eds.), *Woman in sexist society* (pp. 207–224). New York: Basic Books.

Worell, J. (2000). Feminism in psychology: Revolution or evolution? *Annals of the American Academy of Political and Social Science*, *571*, 183–196.

Worell, J., & Johnson, G. N. (1997). *Shaping the future of feminist psychology: Education, research, and practice*. Washington, DC: American Psychological Association.



Forensic Mental Health Assessment

Kirk Heilbrun and Tammy Lander
Drexel University, Philadelphia, Pennsylvania, USA

1. Introduction
2. Context and Principles of FMHA
3. Clinical–Legal Issues
4. Important Developments
5. Future Directions
6. Conclusions
Further Reading

GLOSSARY

forensic mental health assessment The evaluation of defendants or litigants, usually conducted by mental health professionals such as psychologists, psychiatrists, or social workers, that is performed to assist the court in making better informed decisions or to help the attorney present evidence relevant to his or her case.

John Muhammad and Lee Boyd Malvo were charged in the offenses that became known as the "Washington, DC, sniper shootings." Both were charged with capital (death penalty) offenses and tried separately. What forensic mental health assessment might be important in such cases? Possibilities include juvenile transfer to criminal (adult) court, competence to stand trial, competence to proceed as one's own attorney (Muhammad briefly attempted to act as his own attorney), sanity at the time of the offense (Malvo entered a plea of not guilty by reason of insanity), and aggravating and mitigating factors in capital sentencing (both Muhammad and Malvo were convicted of capital murder). This article will help you understand the nature of forensic mental health assessment, the kinds of questions it tries to answer, as well as the limits of its applicability and usefulness.

1. INTRODUCTION

The field of forensic psychology has developed almost entirely during the past 50 years. The broad meaning of "forensic psychology" might include the research or professional activities of any psychologist—developmental, social, cognitive, human experimental, or clinical—when performed in anticipation that the work might be used by a judge or attorney as evidence to promote better informed legal decision making. However, this article focuses more specifically on one aspect of forensic psychology: assessment. This aspect, forensic mental health assessment (FMHA), refers to the evaluation of defendants or litigants by psychologists (or other mental health professionals such as psychiatrists or social workers) that is performed to assist the court in making better informed decisions or to help the attorney present evidence relevant to his or her case.

Described in this way, forensic psychology is a fairly new specialty within the broader field of psychology. Psychologists have not always been welcomed by the courts. In an important 1962 case, *Jenkins v. United States*, a federal appellate court ruled that psychologists with appropriate training and expertise were qualified

©2004 Elsevier Inc.
All rights reserved.

to offer expert testimony on questions regarding mental disorder. Prior to *Jenkins*, courts typically required mental health experts to have a medical degree. Since the *Jenkins* decision, however, there has been a steady increase in psychologists' involvement in legal questions requiring mental health expertise.

Just how frequently do forensic psychologists provide FMHA? We do not have good studies that would answer this question, but the estimated frequency for certain kinds of evaluations, such as criminal defendants' competence to stand trial, is more than 25,000 annually, according to Otto and Heilbrun. Other common legal questions that may be informed by FMHA include competence to waive Miranda rights, sanity at the time of the offense, sentencing (capital and noncapital), juvenile commitment, juvenile waiver to criminal court and reverse waiver back to juvenile court, child custody, and personal injury. Psychologists are involved in conducting many of the evaluations performed in these areas.

2. CONTEXT AND PRINCIPLES OF FMHA

2.1. Sources of Authority in Forensic Assessment

2.1.1. Courts, Statutes, and Case Law

The law is the first place forensic psychologists look in structuring FMHA. Many legal questions are defined, to some extent, by existing statutes. When the criteria for a given legal decision are described in a statute or related evidence code or administrative code, these criteria can be used in the report and possible testimony. This use indicates that the psychologist is both familiar with the relevant law and has structured the evaluation around it, thus making the evaluation more relevant to the proceedings. For any given issue, it is up to the court to answer the ultimate legal question, so it is not the most desirable focus for the report. However, it is helpful to show that the evaluator is aware of the legal standard and can use the defining criteria to structure the report.

Other sources of legal authority may also be helpful, although they may be less directly relevant to the legal criteria used in the report. Appellate case law (particularly cases from the U.S. Supreme Court) and "model" law (e.g., the *Criminal Justice Mental Health Standards*, published in 1989 by the American Bar Association) can provide useful background on the sources of and reasoning behind the current legal standard in the jurisdiction in which the evaluation is being conducted.

2.1.2. Ethical Standards

Forensic psychologists rely on two major sources of ethical authority. First, the *Ethical Principles of Psychologists and Code of Conduct* (Ethics Code), published by the American Psychological Association (APA) in 2002, serves as the strongest authority for ethics in psychology. Those conducting FMHA are ethically obligated to do so in a way consistent with the ethics code and can be sanctioned by the APA for violating it. A second important source of guidance is the *Specialty Guidelines for Forensic Psychologists*, published by the Committee on Ethical Guidelines for Forensic Psychologists in 1991. This document was written for those who regularly conduct FMHA, and it provides more specific guidance for those in legal settings. Finally, and most specifically, there are documents such as the "Guidelines for Child Custody Evaluations in Divorce Proceedings," published by the APA in 1994, that offer direction for particular kinds of FMHA such as child custody evaluation.

2.1.3. Scientific Literature

FMHA is based on and informed by science. Studies in areas such as mental health disorders, behavioral functioning, violence and antisocial behavior, response to treatment, and the assessment of specific functional legal capacities are all important to the development of FMHA procedures and the practice of FMHA in particular cases. Under the 1993 holding of the U.S. Supreme Court in *Daubert vs. Merrell Dow Pharmaceuticals, Inc.*, and the subsequent 1999 Supreme Court decision in *Kumho Tire Company, Ltd. vs. Carmichael*, FMHA evidence is weighed by the trial court judge for admissibility based in part on whether it is based on procedures consistent with science. (Not all states have the *Daubert* standard for admissibility of expert evidence; some continue to use the standard described in the 1923 *Frye vs. United States* case, in which a technique must be sufficiently established to have gained general acceptance in the field to which it belongs to yield admissible expert evidence.)

2.1.4. Professional Literature

Finally, there is a literature that might best be described as "professional," describing recommended approaches to FMHA based on experience and emerging or

established standards in the field. Often such literature appears in practice-oriented journals or books, particularly those that are interdisciplinary. Leading interdisciplinary journals in forensic psychology include *Law and Human Behavior*, *Behavioral Sciences & the Law*, *Criminal Justice and Behavior*, *International Journal of Forensic Mental Health Services*, and others focusing both on empirical studies and on more practice- or theory-oriented approaches to FMHA. This literature forms another important source of authority and guidance for the forensic psychologist.

2.2. Principles of FMHA

Evaluations conducted by psychologists are typically performed to help reach a diagnosis or problem formulation and to plan appropriate treatment (we might call these evaluations "therapeutic assessment"). However, FMHA differs from this usual process and goals in some important ways. Heilbrun pointed out the following 10 differences (Table I) in his 2001 book, *Principles of Forensic Mental Health Assessment*:

TABLE I
Principles of Forensic Mental Health Assessment[a]

Preparation
1. Identify relevant forensic issues.
2. Accept referrals only within area of expertise.
3. Decline the referral when evaluator impartiality is unlikely.
4. Clarify the evaluator's role with the attorney.
5. Clarify financial arrangements.
6. Obtain appropriate authorization.
7. Avoid playing the dual role of therapist and forensic evaluator.
8. Determine the particular role to be played within the forensic assessment if the referral is accepted.
9. Select the most appropriate model to guide data gathering, interpretation, and communication.

Data collection
10. Use multiple sources of information for each area being assessed.
11. Use relevance and reliability (validity) as guides for seeking information and selecting data sources.
12. Obtain relevant historical information.
13. Assess clinical characteristics in relevant, reliable, and valid ways.
14. Assess legally relevant behavior.
15. Ensure that conditions for evaluation are quiet, private, and distraction free.
16. Provide appropriate notification of purpose and/or obtain appropriate authorization before beginning.
17. Determine whether the individual understands the purpose of the evaluation and the associated limits on confidentiality.

Data interpretation
18. Use third-party information is assessing response style.
19. Use testing when indicated in assessing response style.
20. Use case-specific (idiographic) evidence in assessing clinical condition, functional abilities, and causal connection.
21. Use nomothetic evidence in assessing clinical condition, functional abilities, and causal connection.
22. Use scientific reasoning in assessing causal connection between clinical condition and functional abilities.
23. Do not answer the ultimate legal question.
24. Describe findings and limits so that they need change little under cross-examination.

Communication
25. Attribute information to sources.
26. Use plain language; avoid technical jargon.
27. Write report in sections, according to model and procedures.
28. Base testimony on the results of the properly performed FMHA.
29. Testify effectively.

[a]Adapted from Heilbrun (2001).

1. The purpose for which the evaluation is conducted
2. The nature of the relationship between the evaluator and the individual being evaluated
3. The need for a formal notification of purpose
4. Who is being served
5. The nature of the standard(s) applied in the evaluation
6. The sources of data considered
7. The response style of the individual being evaluated
8. The need to clarify the reasoning and limits of knowledge
9. The nature of the written report
10. The expectation of expert testimony

When these differences are considered, it becomes very clear that FMHA is not simply psychological assessment that is conducted in a legal setting. Instead, it is a specialized form of psychological assessment that takes into account that there may be very different questions asked, standards applied, motivations observed, and approaches to communicating results.

General principles that describe therapeutic assessment do not necessarily apply to FMHA, so it is helpful to review the FMHA principles described by Heilbrun in 2001. These principles, discussed in the following section, are summarized and elaborated in how they might apply in FMHA in Table II.

These 29 FMHA principles each apply, to some extent, to the many forms of FMHA. Heilbrun *et al*. provided examples of case reports on different legal questions to which these principles can be applied, so it seems that both theoretically and practically these principles are potentially useful. They are reviewed in the remainder of this section, with a comment about each principle and the associated general guideline:

1. *Identify relevant forensic issues*. This principle distinguishes between the legal question, to be decided by the judge or the jury in the case, and the forensic issues, which are the capacities and abilities included within the legal question. These capacities and abilities can be described and measured more readily than the broader legal question, which typically contains questions of both science (e.g., Did the defendant have a mental illness?) and values (e.g., Was the individual sufficiently disturbed to be unable to know the wrongfulness of his behavior?). The associated guideline suggests that both the legal question and the forensic issues be described in the first section of the report.

2. *Accept referrals only within area of expertise*. Expertise in FMHA has two components: (i) clinical and didactic training and experience with individuals similar to the one being evaluated and (ii) forensic training and experience in evaluating similar individuals and questions. The forensic psychologist can demonstrate both kinds of expertise by submitting a professional resume (a *curriculum vitae*) when requested by the lawyers or the judge and by displaying professional credentials, such as degree, licensure, and board certification, on the letterhead used for the report.

3. *Decline the referral when evaluator impartiality is unlikely*. The evaluator's impartiality is important in FMHA. If the psychologist has personal beliefs or experiences circumstances that would interfere with his or her ability to be fair and even-handed in collecting data and drawing conclusions, then this principle and guideline suggest that the forensic clinician should decide not to conduct the FMHA in this particular case. For example, psychologists with strong feelings about the death penalty that they cannot set aside in a particular case should decline to conduct a capital sentencing evaluation.

4. *Clarify the evaluator's role with the attorney*. When psychologists conduct an FMHA, they may do so upon order of the judge or by request of one of the attorneys involved in the litigation. A psychologist may also serve as a consultant to one of the attorneys, without any expectation that a report or expert testimony will result. These roles have differing expectations for who will receive the results of the evaluation or consultation and in what circumstances. These expectations stem from relevant evidence law, applicable ethical guidelines, and professional practice standards. It is important, though, that the psychologist make it clear which role is being played so the appropriate expectations will guide his or her professional conduct. When roles are unclear or merged, there can be substantial problems with the evaluation or consultation.

5. *Clarify financial arrangements*. Psychologists are paid for conducting FMHA as they would be for any professional service. The details of such payment, however—the rate or total amount, who will be responsible for payment, and whether there are special arrangements to be made (such as a retainer or an escrow account)—should be agreed upon by all parties before the FMHA is begun. If the payment process is established under law or policy, then arrangements are not needed; the evaluator simply bills the court, the government, or whoever is designated under policy for the purpose. In other instances, however, the financial arrangements must be discussed and clarified in the beginning.

TABLE II
Principles and Guidelines for Application of FMHA Principles in Forensic Contexts[a]

Principle	*General guidelines for application in FMHA*
Identify relevant forensic issues.	Cite legal question and include forensic issues in first section of report.
Accept referrals only within area of expertise.	Give degree and licensure, board certification status. Provide curriculum vitae and/or summary of qualifications if requested.
Decline the referral when evaluator impartiality is unlikely.	Avoid involvement in cases in which there is substantial incentive for the forensic clinician (personal, professional, or monetary) to have case decided in a particular direction.
Clarify the evaluator's role with the attorney.	Ensure that both the forensic clinician and the referral source are clear on whether the clinician will serve as court-appointed evaluator, attorney-requested evaluator, or consultant.
Clarify financial arrangements.	Ensure that terms of payment for evaluation are understood by both the forensic clinician and the party responsible for payment.
Obtain appropriate authorization.	Cite basis for evaluation request (e.g., court ordered or attorney requested). Describe whether informed consent was obtained, if evaluation was not court ordered.
Avoid playing the dual roles of therapist and forensic evaluator.	Minimize the frequency of this combination. If such roles are combined, it should be with explicit justification, advance planning, and clear notification to the individual involved.
Determine the particular role to be played within forensic assessment if the referral is accepted.	If report will be submitted into evidence, evaluator should be *impartial*—tone of report should reflect this.
Select the most appropriate model to guide data gathering, interpretation, and communication.	Use the Morse model (mental disorder, functional abilities, and causal connection) or the Grisso model (functional, contextual, causal, interactive, judgmental, and dispositional characteristics).
Use multiple sources of information for each area being assessed.	Obtain self-report, psychological testing data, third-party interviews, and collateral records data.
Use relevance and reliability (validity) as guides for seeking information and selecting data sources.	Use data sources with demonstrated reliability and validity (when this has been researched) and that will provide information relevant to the area being assessed.
Obtain relevant historical information.	In a separate section, document the individual's history and previous functioning in areas relevant to current clinical condition and functional legal capacities.
Assess clinical characteristics in relevant, reliable, and valid ways.	Describe clinical characteristics using measures that are reliable, valid for the purpose used, and/or weighed against information from collateral sources.
Assess legally relevant behavior.	Document information collected from multiple sources regarding the individual's functional legal capacities.
Ensure that conditions for evaluation are quiet, private, and distraction free.	Note any deviation from reasonably quiet, private, and distraction-free conditions. Describe impact on data collected.
Provide appropriate notification of purpose and/or obtain appropriate authorization before beginning.	Describe elements of notification of purpose or informed consent given to individual being evaluated and to third parties who are interviewed.

Continues

Continued

Principle	*General guidelines for application in FMHA*
Determine whether the individual understands the purpose of the evaluation and the associated limits on confidentiality.	Document how the individual's understanding was assessed and to what extent he or she understood the relevant information.
Use third-party information in assessing response style.	Describe the consistency of third-party information with self-reported information, and be particular cautious about self-report when it is significantly different from third-party accounts.
Use testing when indicated in assessing response style.	Administera test(s) sensitive to response style, particularly when there is concern about the accuracy of self-report.
Use case-specific (idiographic) evidence in assessing clinical condition, functional abilities, and causal connection.	Describe the individual's clinical condition and functional legal abilities in the context of his or her history of symptoms and demonstrated capacities.
Use nomothetic evidence in assessing clinical condition, functional abilities, and causal connection.	Describe the results of psychological tests, structured instruments, and specialized tools validated for assessing (i) clinical condition or (ii) functional legal capacities.
Use scientific reasoning in assessing causal connection between clinical condition and functional abilities.	Describe explanations for clinical condition and functional abilities that have the most supporting evidence and least disconfirming evidence.
	When evidence is mixed, or competing explanations seem comparably well-supported, say so.
Do not answer the ultimate legal question.	Present conclusions about forensic capacities but not the larger legal question(s).
Describe findings and limits so that they need change little under cross-examination.	Be careful, impartial, and thorough in presenting data and reasoning.
	Consider alternative explanations.
Attribute information to sources.	Describe data so that the source(s) of any specific finding is clear.
Use plain language; avoid technical jargon.	Make minimal use of technical language, and define technical terms when they must be used.
Write report in sections, according to model and procedures.	Include sections on referral information, sources of information, relevant history, clinical functioning, relevant functional legal capacities, and conclusions.
	Describe causal relationship between clinical symptoms and functional legal capacities.
Base testimony on the results of the properly performed FMHA.	Master the contents of the report, which contains thorough documentation of evaluation, and use report contents to guide testimony.
Testify effectively.	Use effective style in presenting substantive FMHA findings in testimony.

[a]Adapted from Heilbrun (2001).

6. *Obtain appropriate authorization*. The nature of the authorization needed to conduct the evaluation varies, depending on who requested it and the role being played by the evaluator. The forensic clinician needs a signed order from the court for a court-ordered evaluation. Consent from both the referring attorney and the individual being evaluated is necessary when the forensic clinician is in the role of expert for the defense/prosecution/plaintiff.

7. *Avoid playing the dual roles of therapist and forensic evaluator*. Serving as both therapist and forensic evaluator for the same individual should be avoided because it can create significant problems for the psychologist, the person being evaluated and treated, the

lawyers, and the judge. Occasionally, such a dual role cannot be avoided. When it cannot, the guideline emphasizes the importance of making the reason clear, doing the planning in advance, and informing the affected individual(s) in plain language.

8. *Determine the particular role to be played within forensic assessment if the referral is accepted*. If the forensic psychologist respects the previous principle, then most FMHA cases will involve only a single role. The psychologist must identify this role first. When it involves writing a report and possibly testifying, the psychologist must strive to be impartial, focusing only on the questions that are requested. In addition, this role should be maintained throughout the entire case.

9. *Select the most appropriate model to guide data gathering, interpretation, and communication*. Stephen Morse, in 1978, and Thomas Grisso, in 1986, each described a model to guide the ideas, procedures, interpretation of results, and reasoning that are important components of FMHA. Each model helps the psychologist describe the relationship between clinical symptoms and functional legal capacities, making it easier for the reader to understand how the individual being evaluated might meet the applicable legal test. They also help describe the deficits that keep him or her from doing so and the source of these deficits when the individual does not.

10. *Use multiple sources of information for each area being assessed*. A single source of information is usually less accurate than multiple sources considered together. This is particularly applicable in FMHA because some of the information sources can be particularly distorted. A defendant might be inclined to exaggerate or minimize symptoms or capacities, depending on the possible implications of each in litigation. A psychological test might not be sensitive to such deliberate exaggeration or minimization. An individual who was interviewed to provide more information about the defendant might have some bias against or for the defendant. When multiple sources are considered, however, the psychologist can be more sensitive to particular distortions from any single source and better gauge the overall accuracy through considering consistency across sources.

11. *Use relevance and reliability (validity) as guides for seeking information and selecting data sources*. Whether evidence will be admitted into a legal proceeding depends on its relevance to the matter before the court and its reliability (accuracy). Therefore, this principle emphasizes both relevance and reliability in sources of information that are part of FMHA. Some sources of information, such as psychological tests and structured interviews, have measured levels of reliability and validity as part of their development. It is fairly straightforward to select such tests and interviews with higher levels of reliability and validity. Other sources of information, such as the interviews with the individual being evaluated and with collateral observers, may be more difficult to gauge in their validity—but may be selected because they are so clearly relevant to the issues being assessed.

12. *Obtain relevant historical information*. Relevant historical information about the individual being evaluated is almost always important in FMHA. The level of detail needed varies, however, according to the type of FMHA. Some evaluations, such as those of competence to stand trial, are present-state oriented and need a more focused history in areas such as criminal and mental health history. Other evaluations, such as capital sentencing, are much broader and demand a history that is much wider in scope.

13. *Assess clinical characteristics in relevant, reliable, and valid ways*. The decision about which symptoms and problems should be assessed, and how to assess them, is informed by using relevance and reliability, according to this principle. This guideline calls for using appropriate measures, such as psychological tests and structured interviews, and considering these results in conjunction with information from collateral sources in order to assess the relevant clinical aspects under consideration.

14. *Assess legally relevant behavior*. There are particular capacities and behaviors that are part of the specific legal question(s) being considered by the court. These capacities and behaviors (sometimes called functional legal criteria) are a very important part of FMHA. For some kinds of evaluations, there are validated measures that can be used to help the evaluator learn how the individual being evaluated compares with other such individuals on certain relevant legal capacities. Such measures have been developed for competence to stand trial, competence to consent to treatment, and competence to waive Miranda rights. However, for many legal questions, there has not yet been a validated measure developed. Therefore, evaluators must identify the functional legal demands and obtain information about the individual's thinking, feeling, and behavior at the appropriate time, when they do not have a validated measure available.

15. *Ensure that conditions for evaluation are quiet, private, and distraction free*. FMHA is often conducted in detention centers, jails, prisons, or secure

psychiatric hospitals. Security is important in these settings; staff are responsible for preventing escape, harm to others, or harm to self. The need for evaluation conditions that are quiet and distraction free (without significant noise or visual distraction) as well as private (in which communication cannot be overheard by a third party) must therefore compete with security, sometimes in the context of a physical plant that was not designed to facilitate privacy or quiet. This principle emphasizes that the FMHA cannot be performed meaningfully unless there is a certain degree of privacy and limit on distraction afforded during the evaluation period.

16. *Provide appropriate notification of purpose and/or obtain appropriate authorization before beginning*. The forensic psychologist provides a notification of purpose or obtains informed consent before starting the evaluation. Informed consent is needed when an attorney retains a forensic psychologist to evaluate the attorney's client because the evaluation is legally voluntary and the client is under no legal obligation to participate. However, a client undergoing a court-ordered evaluation does not have a legal right to refuse to participate. An evaluee in these circumstances should receive the same information that would be provided as part of obtaining informed consent—the nature and purpose of the evaluation, its authorization by the court, and the associated limits on confidentiality, including how it might be used. However, since there is no legal right to refuse, the necessity for consent is eliminated when a court-ordered evaluation is conducted.

17. *Determine whether the individual understands the purpose of the evaluation and associated limits on confidentiality*. If the information provided in the notification of purpose or request for informed consent is not understood, or is only partly understood, then the process is less meaningful. The forensic psychologist must gauge how well such information is understood by the individual being evaluated. To the extent it is not understood, the psychologist should describe the limits on understanding and what deficits interfere with such understanding.

18. *Use third-party information in assessing response style*. Some individuals evaluated through FMHA respond to questions openly and honestly. Others exaggerate or fabricate symptoms and problems. Yet others minimize or deny symptoms that are genuinely experienced. Response style is a term used to describe how an individual reports his or her own symptoms, behaviors, and problems, and it is used to classify an individual's responding as (i) reliable, (ii) exaggerated/malingering, (iii) defensive, or (iv) uncooperative. Self-reported information from an individual who is not reliable will have significant inaccuracies. This principle emphasizes the value of third-party information (records and interviews of collateral observers) in both detecting and correcting these inaccuracies.

19. *Use testing when indicated in assessing response style*. In addition to third-party information, there are also tests that have been developed and validated to assess response style. These are reviewed by Rogers in his 1997 book titled *Clinical Assessment of Malingering and Dissimulation*. This principle advocates using tests and specialized measures whenever available and appropriate for assessing malingering or defensiveness.

20. *Use case-specific (idiographic) evidence in assessing clinical condition, functional abilities, and causal connection*. Science is an important part of FMHA. The first way it can contribute is through promoting the gathering of individualized information, specific to the case and present functioning of the individual, and comparing it to that individual's capacities and functioning at other times. This approach is consistent with the law's goal of individualized justice, so it may appeal particularly to judges and attorneys.

21. *Use nomothetic evidence in assessing clinical condition, functional abilities, and causal connection*. Science can also contribute to FMHA through empirical data collected with groups of individuals similar to the person being evaluated and by using tests developed and validated on similar groups. Such tests facilitate the reliable measurement of relevant abilities and symptoms and allow their comparison with those in "known groups." For example, a validated test to measure capacities relevant to competence to stand trial would allow the psychologist to compare an individual's performance on a given capacity (e.g., understanding the adversarial nature of the legal system) to that of other known groups, such as (i) defendants who had been adjudicated incompetent to stand trial, (ii) defendants for whom the issue had been raised but were not adjudicated incompetent, and (iii) defendants for whom the issue had never been raised.

22. *Use scientific reasoning in assessing causal connection between clinical condition and functional abilities*. Scientific reasoning, as contrasted with scientific data, can also contribute to FMHA. The results obtained from one source of information (such as the interview) can be treated as "hypotheses" to be supported or

rejected through further information obtained from other sources of information. Hypotheses that account for the most information with the simplest explanation are likely to be the most accurate.

23. *Do not answer the ultimate legal question directly*. Many in the field of forensic psychology have debated whether experts should directly answer the "ultimate legal question" before the court. Some have argued that this is permitted under the law of evidence in most jurisdictions, that many judges seem to want it, and that there is little harm in doing so. Others have suggested that ultimate legal questions (such as "Is the defendant competent to stand trial?") have embedded moral and community values to a greater extent than do functional legal capacities (such as whether the defendant understands his charges and can assist in his defense). Thus, answering ultimate legal questions directly intertwines clinical and scientific conclusions with moral values, possibly to the confusion of the judge or jury. This principle suggests that functional capacities rather than ultimate legal questions should be the focus of FMHA.

24. *Describe findings and limits so that they need change little under cross-examination*. FMHA should be a careful and thorough process, incorporating multiple sources of information and accompanied by acknowledgment of the limits on data accuracy and consistency. Using this approach, the forensic clinician can expect that the description of findings and conclusions will not change significantly during cross-examination.

25. *Attribute information to sources*. All information in the FMHA report should be attributed by a specific source(s). This allows the evaluator to describe consistency across sources. It also means that the judge and opposing counsel can directly assess the credibility of specific information by source.

26. *Use plain language; avoid technical jargon*. The "consumers" of FMHA, such as judges, attorneys, and jurors, usually do not have advanced training in psychology or other behavioral science. Thus, it is useful for FMHA results to be communicated without the use of technical jargon. When important points or concepts are conveyed in nontechnical language, it is usually possible to capture their essential meaning. Occasionally, a technical term may be needed, but a parenthetical definition can help ensure that it is understood by the reader or listener in the way its meaning is intended.

27. *Write reports in sections, according to model and procedures*. There is a great deal of information in an FMHA report. Without an organizing structure for the report, the reader can become lost or fail to understand some of the important aspects of the report. One organizing structure for reports was suggested by Heilbrun in his 2001 book *Principles of Forensic Mental Health Assessment*:

1. Referral: with identifying information concerning the individual, his or her characteristics, the nature of the evaluation, and by whom it was requested or ordered
2. Procedures: times and dates of the evaluation, tests or procedures used, different records reviewed, and third-party interviews conducted, as well as documentation of the notification of purpose or informed consent and the degree to which the information was apparently understood
3. Relevant History: containing information from multiple sources describing areas important to the evaluation
4. Current Clinical Condition: appearance, mood, behavior, intellectual functioning, thought, and personality
5. Forensic Capacities: varying according to the nature of the legal questions
6. Conclusions and Recommendations: addressed toward the relevant capacities rather than the ultimate legal questions

28. *Base testimony on the results of the properly performed FMHA*. Contrary to what is often perceived, most of FMHA does not result in expert testimony. Often, reports are accepted and presented in evidence without such testimony. Sometimes, reports may not be used, depending on the circumstances involved in requesting the report (e.g., court ordered vs attorney requested) and whether the attorney has that choice under the evidence law in the jurisdiction. However, when expert testimony is provided, it should be firmly based on the results of a thorough evaluation documented in the FMHA report. This allows the referring attorney to present the expert's findings effectively, the opposing attorney to challenge them, the judge or jury to understand them, and the expert to communicate them.

29. *Testify effectively*. Effective expert testimony combines strong substance with useful style. The substance of FMHA has been discussed in the previous 28 principles. The style of expert testimony involves appearing knowledgeable, credible, likeable, interesting, and professional. Without both substance and style, the expert testimony provided by a forensic psychologist cannot be maximally effective. Strategies and suggestions for improving style while testifying have been documented by Brodsky in his 1991 book, *Testifying in Court:*

Guidelines and Maxims for the Expert Witness, and his 1999 book, *The Expert Expert Witness: More Maxims and Guidelines for Testifying in Court*.

3. CLINICAL–LEGAL ISSUES

3.1. Criminal

Forensic psychologists may provide FMHA that is focused on a variety of legal questions in criminal, civil, and juvenile/family areas. These are described in detail in the 1997 book by Melton *et al*. titled *Psychological Evaluations for the Courts*. In some cases, only one of these questions will be addressed in a single FMHA; in other instances, the forensic psychologist may be asked to address two or three questions in a single evaluation. Some of the most commonly evaluated legal questions evaluated through FMHA for adult criminal defendants include the following:

1. Competence to waive Miranda rights—whether the defendant waived his or her Fifth Amendment right to remain silent and Sixth Amendment right to counsel in a knowing, voluntary, and intelligent fashion when questioned by the police while in custody
2. Competence to stand trial—whether the defendant understands his or her legal situation and can assist counsel in his or her own defense
3. Sanity at the time of the offense—the particular test varies across jurisdictions; in many states, it involves whether the defendant did not know the nature or wrongfulness of criminal behavior as a result of mental disease or defect; in others, there is the additional question of whether the defendant was unable to conform his or her conduct to the requirements of the law as a result of mental disease or defect
4. Non-capital sentencing—often involving the broad questions of the defendant's risk for recidivism and rehabilitation needs and amenability; in specific kinds of cases, such as those involving sexual offenders, the questions are focused particularly on the need for specialized treatment
5. Capital sentencing—whether the defendant has characteristics, symptoms, or history specified in the aggravating and mitigating factors outlined in the law of the jurisdiction

3.2. Civil

In addition to FMHA in criminal contexts, forensic psychologists conduct assessments for judges and attorneys on questions within civil law. Two of the most frequent legal questions for FMHA in the civil area concern (i) civil commitment (whether an individual, because of mental illness, meets the criteria for involuntary hospitalization as presenting a danger to self or others) and (ii) personal injury [often whether an individual, allegedly the victim of a tort (or noncriminal wrong), suffered psychological injury as a consequence, and what might be needed to treat this injury].

3.3. Juvenile/Family

There are several legal questions within the juvenile and family area that are often addressed through FMHA. These include (i) juvenile commitment (whether a youth adjudicated as delinquent should be committed for residential placement or can be served in a less restrictive intermediate placement such as boot camp or in a community-based placement; often involving the criteria of reoffense risk and treatment/rehabilitation needs and amenability), (ii) juvenile transfer and reverse transfer (whether an adolescent charged with a crime should be handled in juvenile court or adult criminal court, with the criteria for assessment often similar to those for juvenile commitment), and (iii) child custody (in divorce cases, the question of whether custody should be awarded solely to one of the divorcing parents or jointly can be informed by FMHA addressing the needs of the children and the parenting capacities of each parent).

4. IMPORTANT DEVELOPMENTS

As forensic psychology has matured, there have been important developments in the areas of training, research, and practice that may enhance the quality of FMHA that is conducted by forensic psychologists.

4.1. Training

Specialized training is crucial to the goal of high-quality forensic practice. One important development in this area during the past decade has been the approval of forensic psychology as a formal applied specialization within the APA. Forensic psychology joins the areas of clinical, counseling, school, behavioral, clinical health, clinical neuropsychology, clinical child, family psychology, and psychoanalysis in psychology as approved APA specializations. This will promote the development and structure of formal postdoctoral training in forensic

psychology because postdoctoral programs will be able to apply for APA accreditation with a specialization in forensic psychology.

There are a number of such postdoctoral fellowship programs in forensic psychology today, which is another important development in the training area. With a total of 16 fellowships available nationally, those training to become forensic psychologists have the opportunity to receive broad-based training in clinical or counseling psychology prior to the doctoral degree followed by specialized forensic training during the fellowship year. Accreditation of such fellowships should further enhance the quality and uniformity of training in FMHA.

However, not all psychologists wishing to provide FMHA have had the opportunity to receive such specialized training. For those who received the doctoral degree before 1980, postdoctoral fellowship training was not available, nor was formal training during internship or doctoral training available except in rare instances. The role of continuing education in the form of workshops presented to practicing professionals has therefore been quite important. Some continuing education is provided by university–state partnerships, as in the state of Virginia. Other continuing education programs are offered by specialty organizations, such as the American Academy of Forensic Psychology. Such continuing education efforts have filled a gap between formal education and experience that was important to more senior psychologists conducting FMHA but have also been helpful in providing relevant research and new tools and approaches to even those who are already largely specialized in forensic.

4.2. Research

The development of specialty tools has been very important in the maturation of the field. Such tools, developed and validated through scientific research and designed to measure functional legal capacities described in the law, have allowed forensic psychologists to consider individualized impressions and compare the individual being evaluated to others in similar circumstances. Tools have been developed during the past decade to measure constructs such as those relevant to competence to waive Miranda rights, competence to stand trial, competence to consent to treatment, violence risk assessment, and response style. As a consequence, FMHA on these issues can be conducted with more sophistication and scientific support than was possible before the development of these specialized, validated FMHA tools.

Foundation support of research networks has been a second significant development in research relevant to FMHA. In particular, the John D. and Catherine T. MacArthur Foundation sponsored the MacArthur Research Network on Mental Health and Law, which developed and carried out important programmatic research in the areas of violence risk, competencies, and coercion in the decade between the late 1980s and the late 1990s. Currently, the MacArthur Foundation sponsors two additional research networks—one on mandated community treatment and the second on juveniles—that have also designed and conducted important research that will enhance the quality and scientific foundation of FMHA conducted in these areas in the future.

One of the areas that have witnessed the most active research and development during the past decade is risk assessment of violent behavior. The field has moved from the relatively primitive approach to violence risk seen in the 1970s and early 1980s to a far more sophisticated, theoretically coherent, and empirically supported process. The research conducted by the MacArthur Research Network on Mental Health and Law is described in the 2001 book by Monahan and colleagues titled *Rethinking Risk Assessment: The MacArthur Study of Mental Disorder and Violence*. In addition to this research, other works have yielded a number of tools assessing both risk factors for violence and needs for risk reduction (sometimes called risk-needs tools). As a consequence of this research, practitioners of FMHA can now address violence risk and risk reduction needs in those they evaluate in a way that is far more coherent and empirically supported than was ever possible before.

4.3. Practice

Finally, there have also been important developments in the practice of forensic psychology, particularly FMHA. Some states (e.g., Massachusetts) have developed training curricula for practicing psychologists in the forensic area and developed a formal certification process for those wishing to become certified examiners in FMHA. More advanced individual certification has been available since the late 1970s from the American Board of Forensic Psychology, which has continued to offer a rigorous board certification examination to those who meet the criteria for training and experience and wish to be certified as forensic practitioners at an advanced level. The development and publication of specialty guidelines for the practice of FMHA was an additional important step. The Committee on Ethical Guidelines for Forensic Psychologists, which published the *Specialty Guidelines*

for Forensic Psychologists in 1991, effectively provided the field with standards that were consistent with the APA ethics code but more specific to forensic contexts, allowing both psychologists and legal professionals a view of aspirational standards of practice in forensic psychology.

5. FUTURE DIRECTIONS

The field of forensic psychology has undergone tremendous change during the past two decades. The practice of FMHA, as an integral part of forensic psychology, has become more legally relevant, better supported by empirical research, and more theoretically sophisticated. There are three developments that are likely to continue this growth and improvement during the next decade. The first involves the development of more formal practice standards and updating the *Forensic Specialty Guidelines* published in 1991 to reflect the changes in the field since that time. The second involves additional research, particularly in the functional legal capacities of legal questions that have not been researched in a way that promotes specialized tool development. With the development of an increased number of empirically supported, legally relevant specialized tools, the overall quality of FMHA will certainly be enhanced. Finally, it will be important to continue to expand formal training in forensic psychology. In part, this will involve the accreditation of postdoctoral fellowships specializing in forensic psychology and the development of additional fellowships. As more psychologists receive forensic specialty training, however, it will become feasible to offer less specialized but still relevant forensic training at the predoctoral and internship levels to psychologists who would provide FMHA less frequently or in a more targeted fashion but who still wish to provide courts and attorneys with a product that promotes better informed legal decision making in areas served by the field of forensic psychology.

6. CONCLUSIONS

How might the questions raised in the Washington, DC, sniper case be addressed through FMHA? One of the defendants, John Muhammad, briefly attempted to act as his own attorney. Was he competent to do so? Although this question is evaluated less frequently than a defendant's competence to stand trial, it is one that could be addressed through FMHA that described his mental state and his abilities under relevant law to represent himself. Since Mr. Muhammad was tried and convicted of capital murder, the jury had to consider whether he should receive a sentence of life in prison or death. FMHA describing Mr. Muhammad in terms of the aggravating and mitigating factors under the law of Virginia could have helped the jury consider this question carefully. Lee Boyd Malvo, Mr. Muhammad's accomplice, was tried separately and also charged with capital murder. Malvo was 17 years old at the time of the offense, so the defense might have argued that he should be tried in juvenile court. (Many states do not permit adolescent defendants charged with murder to be treated as juveniles, but this would have been a possibility if the charge had been less serious.) The defense next argued that Mr. Malvo was insane at the time of the alleged offense. Both defense and prosecution introduced evidence based on FMHA conducted by various experts describing Malvo's mental state at the time of the offense. Finally, FMHA on capital sentencing, describing the existing aggravating and mitigating factors for this defendant, could (and did) help inform the jury considering the death sentence. The Washington sniper case illustrates the different ways that FMHA could be used to assist the court in making various kinds of decisions. The continued growth and expansion of the field of forensic psychology promises to bring more relevant and reliable information into the courtroom to continue to assist legal decision makers.

See Also the Following Articles

Legal Competency ■ Psychology and the Law, Overview

Further Reading

Brodsky, S. (1991). *Testifying in court: Guidelines and maxims for the expert witness*. Washington, DC: American Psychological Association.

Brodsky, S. (1999). *The expert expert witness: More maxims and guidelines for testifying in court*. Washington, DC: American Psychological Association.

Committee on Ethical Guidelines for Forensic Psychologists. (1991). Specialty guidelines for forensic psychologists. *Law and Human Behavior*, *15*, 655–665.

Daubert v. Merrell Dow Pharmaceuticals, Inc., 113 S.Ct. 2786 (1993).

Frye v. United States, 293 F. 1013 (D.C. Cir 1923).

Goldstein, A. (Ed.). (2003). Forensic psychology. In I. Weiner (Editor-in-Chief), *Comprehensive handbook of psychology* (Vol. 11). New York: Wiley.

Grisso, T. (1998). *Forensic evaluation of juveniles*. Sarasota, FL: Professional Resource Press.

Grisso, T. (2002). *Evaluating competencies* (2nd ed.). New York: Kluwer/Plenum.

Heilbrun, K. (2001). *Principles of forensic mental health assessment*. New York: Kluwer/Plenum.

Heilbrun, K., Marczyk, G., & DeMatteo, D. (2002). *Forensic mental health assessment: A casebook*. New York: Oxford University Press.

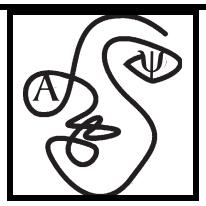
Hess, A., & Weiner, I. (Eds.). (1999). *The handbook of forensic psychology* (2nd ed.). New York: Wiley.

Kumho Tire Company, Ltd. v. Carmichael, 526 U.S. 137 (1999).

Melton, G., Petrila, J., Poythress, N., & Slobogin, C. (1997). *Psychological evaluations for the courts: A handbook for mental health professionals and attorneys* (2nd ed.) New York: Guilford.

Monahan, J., Steadman, H., Silver, E., Appelbaum, P., Robbins, P., Mulvey, E., Roth, L., Grisso, T., & Banks, S. (2001). *Rethinking risk assessment: The MacArthur study of mental disorder and violence*. New York: Oxford University Press.

Rogers, R. (Ed.). (1997). *Clinical assessment of malingering and deception* (2nd ed.). New York: Guilford.



Full-Service Schools

Adena B. Meyers and Mark E. Swerdlik
Illinois State University, Normal, Illinois, USA

1. Introduction
2. Rationale for Full-Service Schools
3. Generalist Approach to Service Delivery
4. The Obstacle of Categorical Funding
5. Variations in Organizational Structure and Function
6. Psychologists in Full-Service Schools
7. Implications for Continuing Professional Development of Psychologists
8. What Is the Future of Full-Service Schools?
Further Reading

GLOSSARY

categorical funding A system of funding that allocates resources to separate programs with specialized professionals who address narrow domains of functioning.

developmental psychopathology An area of study focused on understanding how emotional, behavioral, and learning problems develop over time. The influence of interacting risk and protective factors is emphasized.

ecological theory A theoretical perspective that views human behavior and human development in the context of interconnected social systems.

intervention Services delivered to ameliorate identified problems in learning, behavior, health, or mental health. Examples include medication, psychotherapy, and behavior modification.

prevention Services delivered to promote healthy functioning thereby reducing the chances that problems will develop.

In addition to the traditional academic instruction typically provided in public schools, full-service schools offer an array of integrated health and social services to children and their families. This model of service delivery is based on the assumption that children will be successful learners if their social, emotional, and physical needs are met, and that many of these needs can be addressed effectively, conveniently, and efficiently through school-based service delivery.

1. INTRODUCTION

Although schools exist primarily to teach children academic skills, links between educational systems and other services have existed for more than a century in the United States (Fig. 1). Public health, in particular, enjoys a natural partnership with public education. Requiring all children attending school to be immunized, for example, benefits schools by protecting students and staff from many communicable diseases and advances the public health goal of universal immunization. As Joy Dryfoos noted in her 1994 book, *Full Service Schools: A Revolution in Health and Social Services for Children and Adolescents*, this collaboration dates back to 1870 when the New York City Board of Health (founded 4 years earlier in response to a cholera outbreak) began working with the New York City Board of Education to vaccinate school children. At the turn of the century, several cities initiated school nursing

©2004 Elsevier Inc.
All rights reserved.

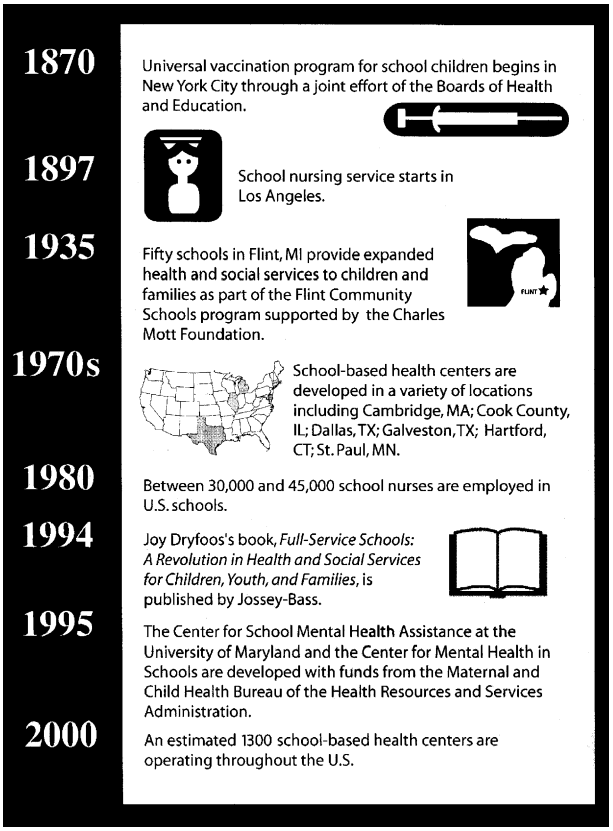

1870
Universal vaccination program for school children begins in New York City through a joint effort of the Boards of Health and Education.
1897
School nursing service starts in Los Angeles.
1935
Fifty schools in Flint, MI provide expanded health and social services to children and families as part of the Flint Community Schools program supported by the Charles Mott Foundation.
FLINT
1970s
School-based health centers are developed in a variety of locations including Cambridge, MA; Cook County, IL; Dallas, TX; Galveston, TX; Hartford, CT; St. Paul, MN.
1980
Between 30,000 and 45,000 school nurses are employed in U.S. schools.
1994
Joy Dryfoos's book, Full-Service Schools: A Revolution in Health and Social Services for Children, Youth, and Families, is published by Jossey-Bass.
1995
The Center for School Mental Health Assistance at the University of Maryland and the Center for Mental Health in Schools are developed with funds from the Maternal and Child Health Bureau of the Health Resources and Services Administration.
2000
An estimated 1300 school-based health centers are operating throughout the U.S.

FIGURE 1 History of expanded school-based services.

programs allowing children to be treated for minor health concerns at school rather than being sent home.

The full-service school model represents an expanded version of such cross-disciplinary collaboration in the provision of services designed to enhance children's social, emotional, physical, and intellectual development. Although the idea has yet to take hold in a permanent, universal way, a number of relevant model programs and social experiments have appeared over the years, with fluctuating support from such sectors as state and federal government, private foundations, and professional organizations such as the American Medical Association.

There is considerable variation in what these programs offer and how they are structured. The label "full-service school" has been applied to schools with on-site health clinics, to schools with expanded hours of operation and social services aimed at violence prevention, and to schools that offer mental health services to students and families. Even schools that do not offer these services on-site may be labeled full-service schools if they have well-developed links (e.g., referral agreements) with relevant agencies in their communities. In this article, the underlying rationale and approach to service delivery that tie these diverse programs together are articulated. Variations in the organizational structure and function of full-service schools, potential roles for psychologists within these service delivery systems, and implications for in-service and preservice training are also discussed.

2. RATIONALE FOR FULL-SERVICE SCHOOLS

2.1. Theoretical Background

Rather than focusing narrowly on academic goals, full-service schools are designed to address children's needs holistically. This approach is consistent with the theoretical perspectives of developmental psychopathology and ecological psychology. Developmental outcomes in the physical, social–emotional, and cognitive domains are assumed to be influenced by a variety of interconnected risk and resilience factors. These include characteristics of the individual child (e.g., temperament and cognitive abilities), the contexts in which the child develops (e.g., family, school, and neighborhood), and transactions within and between these systems. For example, a child's temperament is assumed to influence and be influenced by the skills and parenting style of his or her primary caregivers, which in turn may be shaped by a host of environmental factors, including cultural norms, neighborhood characteristics, and workplace experiences.

2.2. The Developing Child in Context

At the level of the individual child, development may be influenced by biological factors such as prenatal exposure to alcohol, illegal drugs, and tobacco; birth weight; and genetics. When children grow older, their development is also influenced by their own behavior as they engage in (or avoid) such activities as substance abuse, gang involvement, unsafe sex, dangerous driving, and other health-related risky activities.

Children develop in the context of proximal environments or microsystems, including families, schools, neighborhoods, and peer groups, each of which has the potential to influence outcomes. In the family context, the quality and quantity of children's interactions with their primary caregivers often provide a basis for developmental trajectories leading in desirable or undesirable

directions. For many children, these interactions are limited as a result of increasing divorce rates, large numbers of children growing up in single-parent homes, and increasing numbers of two-parent families in which both parents are employed outside of the home. In some cases, family functioning is further compromised by the presence of parents with mental health or substance abuse problems, illiteracy, family violence, or economic strain associated with unemployment and poverty.

At school, children's social and academic experiences may include overcrowded classrooms, bullying, and peer rejection as well as a range of development-enhancing opportunities. In their neighborhoods, many children must contend with risk factors such as poverty, drug trafficking, exposure to traumatic events, and gang activity. Other systemic influences include mass media, which exposes children to images of violence and risky behavior but may also deliver health-promotion messages; government policies and laws, which often dictate how resources are allocated to needy children and families; and many ideologies and values that underlie much of U.S. society. These macrosystemic influences include institutionalized racism and sexism.

Development is influenced not only by the characteristics of these systems but also by the transactions within and between systems. For example, it has been established that risky behavior among adolescents (including violence and unsafe sex) is most likely to occur during late afternoon and early evening hours. To understand this finding, it is necessary not only to consider the role of the family, the school, the neighborhood, or the peer group but also to recognize the ways in which these systems (along with other, more distal systems) work together to increase (or reduce) children's risk levels.

Related to this is the idea that sets of common risk factors (such as those that produce daily blocks of time during which most adolescents within a community have little to do and little adult supervision) often lead to a variety of divergent emotional and behavior problems. It is becoming increasingly clear, for example, that outcomes as conceptually distinct as adolescent pregnancy, school drop-out, violent behavior, and substance use may represent different end points of very similar causal pathways.

2.3. Child Outcomes of Concern

The consequences of the many risk factors described previously are seen across a variety of domains of children's functioning. Specifically, increasing rates of mental health problems, substance use and abuse, and dangerous behavior have been observed among young people in recent years. An estimated 20% of youth between the ages of 9 and 17 have diagnosable mental or addictive disorders, and approximately 4 million children and adolescents suffer from major mental illnesses associated with significant impairment in functioning. Additionally, an increasing number of deaths among young people are attributed to risky or dangerous behavior, including homicide, suicide, HIV-AIDS, and automobile accidents.

2.4. Full-Service Schools: Comprehensive, Integrated, and Preventative

Full-service schools address the risk factors that contribute to these worrisome outcomes by providing comprehensive and integrated educational, mental health, medical, and social services. Prevention programs also represent an important response to the wide variety of biological and environmental risks and stressors affecting school-age children and adolescents. Designed to promote positive quality of life and adaptive functioning across settings, and to reduce the incidence of psychopathology, prevention programs fit naturally into the full-service school model.

One important advantage of full-service schools is the opportunity for families to access coordinated, multidisciplinary services through "one-stop shopping." A second advantage, from both an intervention and prevention perspective, is that full-service schools provide access to most (if not all) children and their families in a key setting. Typically, parents may be more comfortable accessing a variety of mental health, medical, and social services in the more familiar setting of a school as opposed to a local community mental health center.

3. GENERALIST APPROACH TO SERVICE DELIVERY

The full-service school model requires a high level of collaboration among diverse professionals, such as clinical, school, and community psychologists; physicians; nurse practitioners; social workers; family educators; and educators. This integrated, multidisciplinary model represents a generalist approach to service delivery, in which children's functioning and needs are viewed holistically. Whereas traditional service delivery systems divide problems into separate,

specialized domains, each assigned to a different profession (e.g., with medical professionals concerned with physical health, educators concerned with learning, mental health professionals responsible for emotional and behavioral functioning, and social workers responsible for family environment factors), the generalist approach emphasizes the interconnections among these domains.

For example, under a more traditional or specialist model, a child who is having trouble learning to read might be referred to a school psychologist for assessment and a reading specialist or special educator for intensive reading services. The student may have received vision screening through a pediatrician and eyeglasses through a social service agency. Contact between these various agencies and professionals would be limited. The generalist model, in contrast, would simultaneously consider health, mental health, family, and community factors that might influence the child's functioning. Such issues as vision, nutrition, family members' literacy (and English language proficiency), and family resources would be considered as possible explanations for the child's difficulty and would be integrated into any intervention plan.

4. THE OBSTACLE OF CATEGORICAL FUNDING

Traditionally, one of the major barriers to the generalist approach to service delivery is the categorical nature of funding sources. Resources controlled by separate agencies are allocated to separate programs with specialized professionals who address narrow domains of functioning. Proponents of full-service schools have argued that this system results in fragmented, inefficient services that may duplicate efforts, on the one hand, and miss opportunities to influence meaningfully the common factors that underlie many of the devastating outcomes that children experience across domains of functioning, on the other hand. Some foundations as well as federal, state, and local government agencies have responded to this problem by relaxing restrictions on how funds may be used.

5. VARIATIONS IN ORGANIZATIONAL STRUCTURE AND FUNCTION

Full-services schools, also referred to as school-based or school-linked health centers, differ in their degree of comprehensiveness and integration of services, whether services are primarily preventative or treatment oriented, and whether the services are delivered in the school or in locations within the community (Fig. 2). On the surface, it appears that school-based services enjoy certain advantages over school-linked services. These advantages include convenience for the children and their families as well as enhanced opportunities for collaboration between health and

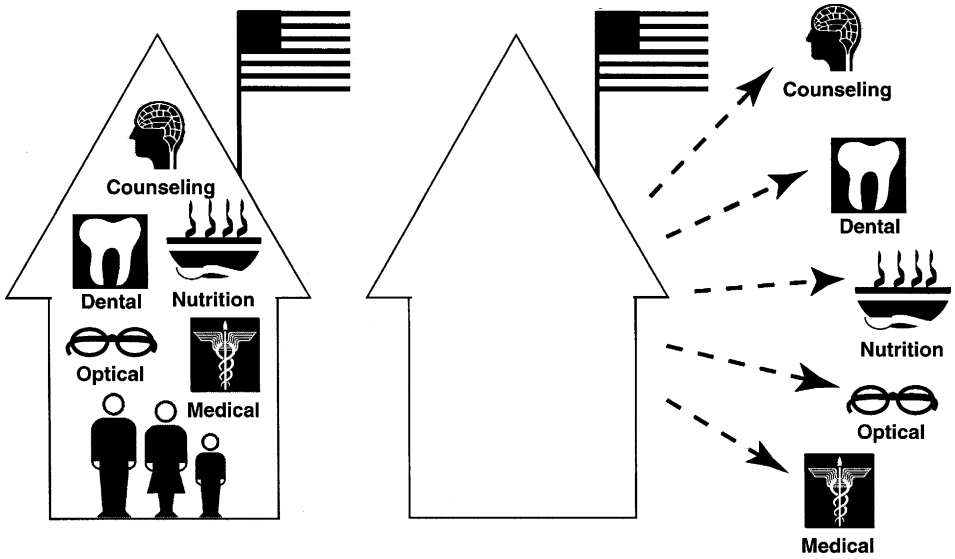


FIGURE 2 School-based vs school-linked models.

social service providers and school personnel, such as teachers, counselors, and administrators.

Despite the intuitive appeal of the distinction between the school-based service models and school-linked models, there is evidence that the location of services may be less important than originally assumed. A nationwide survey conducted by Steven Shaw and colleagues in 1996 found no significant differences between school-based and school-linked services in terms of types of programs offered (acute and preventative medical services, and mental health counseling focused on drugs, alcohol, family, and violence). Neither model provided many social services, such as employment or financial counseling or parole/probation counseling, and both were more likely to offer medical than mental health services. With respect to cross-disciplinary collaboration, it was found that school-based clinics were no more likely to be integrated with education than were school-linked clinics. Thus, despite a large body of literature stressing the importance of a clinic's location as it relates to integration of services, many of the school-based clinics operated autonomously from core educational services.

6. PSYCHOLOGISTS IN FULL-SERVICE SCHOOLS

Psychologists join other professionals in the staffing of full-service schools. Roles for psychologists include providing a range of mental health services, such as screening for conditions such as depression and eating disorders; intervention, including individual, group, and family therapy/counseling; and prevention activities particularly related to substance abuse, pregnancy and sexually transmitted diseases, other high-risk health behaviors, and violence.

Psychologists in full-service schools may also serve as consultants to other professionals, such as teachers, administrators, and medical personnel, with regard to mental health and systemic issues. Psychologists who are skilled in collaborative problem solving and teaming may serve as effective team leaders.

Related to the role of collaborative consultant, psychologists may also function as full-service school administrators and grant writers. Many different funding sources exist for full-service schools, including state and federal mental health, public health, compensatory, and special education funds and also third-party funding such as private insurance and Medicaid. Funding from private sources, including the Robert Wood Johnson and Hogg Foundations, and public and private hospitals also supports the development of full-service schools.

Researcher/program evaluator represents another potential role for the psychologist. As is true in any new initiative, it is important to document outcomes for students and their families. However, this is particularly challenging when prevention is one of the services offered. Psychologists with expertise in applied research and program evaluation could help fulfill this function.

7. IMPLICATIONS FOR CONTINUING PROFESSIONAL DEVELOPMENT OF PSYCHOLOGISTS

Efforts to expand and reconceptualize current training are called for to help prepare psychologists to offer the broad range of services necessary to perform many of the previously mentioned functions in full-service schools. These changes could occur at the preservice level, including practicum and internship experiences, as well as at the in-service level, including opportunities for continuing professional development. Relevant foci for training include the rapidly growing research base on behavioral health and behavioral management of medical problems. In addition, prevention of academic failure, mental health problems, and health risk behaviors such as smoking and substance abuse as well as promoting healthy development through the use of health education, health services, and guidance are important areas of emphasis.

Because learning is one of the primary missions of the school, psychologists would benefit from knowing how factors such as health, nutrition, substance abuse, and conflict with family or peers influence the learning process. Empirically supported interventions to address these factors at the individual, community, and societal level should also be included in preservice and continuing professional development programs for psychologists who wish to work in full-service schools. Also important are such areas as psychopharmacology (e.g., the effects of various medications on children's learning and behavior), family systems, health promotion, and individual and group therapy. Knowledge of interventions that target multiple microsystems of the child's environment,

including the classroom, the family, peers, the school, and the community as a whole, may be required.

Development of full-service schools will involve significant system reform. In order to succeed, the mission of the full-service school must be integrated into the mission of the school and communicated to decision makers. Thus, information about school reform efforts and initiatives should also be provided as part of the preservice and continuing professional development of psychologists who want to work in full-service schools. Understanding how expanded services fit into the typical organization and structure of the school and familiarity with system change strategies are particularly important for the psychologist. Background in community psychology would also be helpful, as would training in program administration and supervision for the role of program administrator.

Training in group dynamics, collaborative problem solving/consultation, and teaming represents a necessary focus for those who wish to work in the collaborative environment of a full-service school. Areas of concentration include systems-level consultation, coordination of service delivery, obtaining financial reimbursement for services, and overcoming resistance.

Training in the use of technology in providing effective health services to students and their families would be beneficial as well. Technology can be used for data management to enhance and support intervention evaluation activities and to empower students and their families to develop self-help skills and increase self-sufficiency. Technology, including the Internet, can be used to access a wealth of information and resources.

8. WHAT IS THE FUTURE OF FULL-SERVICE SCHOOLS?

In the future, schools will likely be challenged to meet the increasingly complex academic, physical, and mental health needs of children and their families. One way to accomplish this may be to expand the network of full-service schools from the current level of approximately 1300 nationally. In order for this movement to reach its objectives, reform efforts are needed to restructure schools beyond instructional and management systems to providing comprehensive, integrated services addressing all barriers to learning and promoting physical and mental health. This includes continuing efforts to make effective coordinated health care available to poor children and children of color.

To achieve the goals of the full-service school, more effective methods of coordinating and integrating existing school and community resources must be developed. Disciplines involved with full-service schools will need to address barriers to more effective coordinated efforts. One way to do this is to provide interdisciplinary training in evidence-based prevention and interventions at the graduate, postgraduate, and continuing education levels to allow health, mental health, and education personnel to work collaboratively and effectively in full-service schools. In this context, psychologists may be able to assume leadership roles in outcomes-based program evaluation activities so as to apply the science of psychology to better inform practice, resulting in implementation of maximally effective full-service schools.

See Also the Following Articles

Behavioral Observation in Schools ■ Learning ■ Learning Styles and Approaches to Studying ■ Neuropsychological Assessment in Schools ■ School–Community Partnerships

Further Reading

Adelman, H. S., & Taylor, L. (1994). *On understanding intervention in psychology and education.* Westport, CT: Praeger.

Adelman, H. S., & Taylor, L. (1998). Reframing mental health in schools and expanding school reform. *Educational Psychologist*, *33*, 135–152.

Adelman, H. S., & Taylor, L. (2000). Shaping the future of mental health in the schools. *Psychology in the Schools*, *37*, 49–60.

Bronfenbrenner, U. (1977). Towards an experimental ecology of human development. *American Psychologist*, *32*, 513–531.

Brown, M. B., & Bolen, L. M. (1993). School-based health centers. [Special issue]. *School Psychology Review*, *40*(2).

Bucy, J. E., Meyers, A. B., & Swerdlik, M. E. (2002). Best practices in working in full service schools. In A. Thomas, & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 281–291). Bethesda, MD: National Association of School Psychologists.

Carlson, C. D., Tharinger, D., Bricklin, P., DeMers, S., & Paavola, J. (1996). Health care reform and psychological practice in the schools. *Professional Psychology: Research and Practice*, *27*, 14–23.

Center for School Mental Health Assistance. http://csmha.umaryland.edu.

Dryfoos, J. (1994). *Full service schools: A revolution in health and social services for children, youth, and families.* San Francisco: Jossey-Bass.

Paavola, J. C., Cobb, C., Illback, R. J., Joseph, H. M., Jr., Torreulla, A., & Talley, R. C. (1995). *Comprehensive and coordinated psychological services for children: A call for service integration*. Washington, DC: American Psychological Association.

Reeder, G. D., Maccow, G. C., Shaw, S. R., Swerdlik, M. E., Horton, C. B., & Foster, P. (1997). School psychologists and full-service schools: Partnerships with medical, mental health, and social services. *School Psychology Review*, *26*, 603–621.

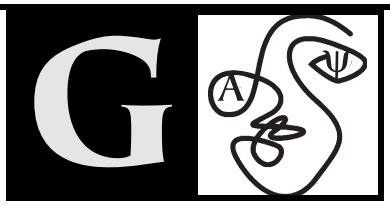
Shaw, S. R., Reeder, G., Swerdlik, M. E., Maccow, G. C., Horton, C., & Bucy, J. (1996, August). *A national survey of full-service schools: Services provided, professionals represented, and the role of the school psychologists*. Paper presented at the annual meeting of the American Psychological Association, Toronto, Canada.

Talley, R. C., Short, R. L., & Kolbe, L. (1995). *School health: Psychology's role: A report to the nation*. Washington, DC: American Psychological Association Center for Psychology in the Schools and Education/U.S. Public Health Service, Centers for Disease Control and Prevention.

UCLA Center for Mental Health in the Schools. http://smhp.psych.ucla.edu.

Weist, M. D. (1999). Improving children's mental health through expanded school-based programs [Special issue]. *Clinical Psychology Review*, *19*(2).

Weist, M. D., Myers, P. C., Danforth, J., McNeil, D. W., Ollendick, T. H., & Hawkins, R. (2000). Expanded school mental health services: Assessing needs related to school level and geography. *Community Mental Health Journal*, *36*, 259–273.



Gender and Culture

Deborah L. Best and Dustin J. Foster
Wake Forest University, Winston-Salem, North Carolina, USA

1. Gender Development across Cultures
2. Behavioral Differences among Boys and Girls
3. Development of Gender Roles and Stereotypes
4. Theories of Gender-Related Learning
5. Gender at the Adult Level
6. Relations between Women and Men
7. Gender Differences at the Cultural Level
8. Conclusion
   Further Reading

GLOSSARY

culture A dynamic system of customary rules and expectations shared by a group of people and incorporating the beliefs, attitudes, norms, social practices, and organizations of that group.

gender Term used to describe the differences between males and females, often by emphasizing social and behavioral factors.

gender roles The social roles (e.g., familial, interpersonal, occupational, and recreational) in which males and females participate with differential frequency.

gender stereotypes The psychological traits that are believed to occur with differential frequency between males and females (e.g., men are more "aggressive" and women are more "emotional").

masculinity/femininity The degree to which females and males perceive that they or others possess traits or behaviors that are considered "woman-like" and "manlike" in their culture.

sex Term generally used to describe the anatomical and physiological differences between males and females, but it may also be related to social and behavioral differences.

sex role ideology Beliefs about the appropriate roles and relationships of men and women. They are often described as varying from traditional to egalitarian.

Gender is a term used to describe a set of psychological, behavioral, and physical attributes that characterize members of either sex. Among any group of people, the psychological and behavioral components of gender often lead to various degrees of maleness and femaleness with considerable overlap between the two. Because gender is defined by the beliefs groups of people have regarding what it means to be male or female, gender is a subjective concept and an appealing topic for psychologists who study multiple cultures. Anyone who has visited different countries has surely noticed differences in the way men and women dress, the kinds of work they do, and the way they interact with one another. Such behavioral differences and their corresponding expectations regarding what is proper for men and women within a particular culture are referred to as gender roles. This article discusses cross-cultural research on gender roles; the degree of cross-gender similarity within cultures in comparison to the substantial variability often observed across cultural groups may be surprising. This article also examines gender theories from developmental,

©2004 Elsevier Inc.
All rights reserved.

social, and personality perspectives, with particular attention to various means of studying and measuring gender across the life span.

1. GENDER DEVELOPMENT ACROSS CULTURES

1.1. Distal Considerations

Developmental influences can be grouped into two somewhat overlapping categories: distal and proximal factors. Distal factors are those experienced by many or all within a culture. These include socioeconomic (e.g., national economy and type/availability of work), political, and religious factors as well as prevalent societal expectations. Discrepancies exist within cultures to the extent that job opportunities, political roles, governmental regulations, and social norms for proper behavior differ for men and women. By growing up with continuous exposure to these overarching factors, children learn the gender role expectations of their culture.

1.1.1. Socioeconomic Influences

National standards of living tend to be related to parents' reasons for having children. The nine-nation Value of Children study found that parents in agrarian cultures (e.g., Indonesia) frequently endorsed utilitarian reasons for having children. These values typically indicated a desire for fulfillment of fundamental needs, such as financial support in old age. In contrast, parents in industrialized nations (e.g., Germany) reported reasons for having children that related to fulfillment of emotional needs, such as affection. Parents in agrarian societies tended to bear more children, perhaps in response to environmental conditions. Particularly relevant to gender development, parents in agrarian societies favored boys over girls, but the two were valued equally by parents in industrialized nations. Similarly, males in societies marked by labor-intensive economies tend to be more successful than females at securing jobs and generating income for their families. Gender role expectations for women may limit their ability to contribute financially to the family, often preventing them from working in jobs outside of the home. Additionally, a young woman may marry into another family and cease contributing to her biological family.

1.1.2. Cultural Environments

Cultures also differ with regard to the religious practices endorsed by the people and/or the political/governmental organization. Religious doctrines vary greatly in the status and freedoms they grant to men and women. The three most predominant religions throughout the world are Christianity, Islam, and Hinduism. In theory, Christian teachings endorse considerable gender equality. However, the Christian God is generally personified as male, and the principal religious servants, ministers, and priests are almost exclusively men. The Hindu pantheon has goddesses as well as gods, and both men and women serve in the temples. In contrast, in Muslim theology, all the significant figures are male and religious practice is controlled by men, as is Muslim society. In nations in which religion and government are mostly separate, religion may not exert a universal influence on the developing child. However, when there is a union between religious doctrine and governmental rule, religious beliefs may strongly influence the relations between men and women as well as child-rearing practices.

Language is one of the most noticeable ways that cultures differ, and indeed this fundamental difference can lead to divergence in gender roles. Languages differ in terms of the frequency and connotation of gender-related words. Recent estimates indicate that approximately 7000 different language systems exist throughout the world. Obviously, the potential for cross-cultural differences in the way children mentally represent and verbally express gender-laden information is substantial. For example, in English it is grammatically correct to use the pronoun "he" to mean he or she, but research has shown that psychologically "he" is not a sex-neutral or generic term and may create a male bias in communication. In inflected languages, such as French or German, which have gender-specific endings for words, it is virtually impossible to communicate without introducing gender. Although such linguistic messages may seem trivial, their prevalence in everyday speech make them a subtle but pervasive influence on sex differentiation practices.

The aforementioned distal environmental influences can interact with biological factors to produce adaptive but distinctive cultural differences. When cross-cultural gender differences exist, they may in part be the result of contrasting environmental triggers that elicit different behaviors. Consistent with this perspective, David Gilmore proposed that the male macho behavior pattern could be an adaptation to divergent economic conditions. This speculation is supported by an

examination of the gender roles of two South Pacific islands, Truk and Tahiti. In the Trukese culture, men are violent, competitive, and sexually promiscuous, whereas women are expected to be submissive. Tahitian culture differs dramatically in that the men engage in submissive behaviors and avoid violent or competitive encounters. In contrast, Tahitian women tend to be sexually active. Gilmore explained these gender role discrepancies by a consideration of each culture's means of obtaining food. Tahitians fish in a protected lagoon where fish are plentiful. However, the Trukese must fish in the open ocean with continuous exposure to life-threatening situations. From this account it appears that the Trukese male gender role is an adaptive response to increase one's chances of survival. This evolutionary theory of gender role development is generally consistent with widely held views of the interaction of nature and nurture. However, many criticisms of the theory exist and research has often failed to support some of its assumptions.

Considering these distal influences, it is undeniable that culture can have profound effects on individual development. Sex differentiation messages are clear and cultural practices that emanate from these determine how babies are delivered, how children are socialized, how women and men dress, what tasks they pursue, what languages they speak, and what roles they adopt. Distal cultural factors are also related to the more immediate, proximal factors in development.

1.2. Proximal Considerations

Proximal developmental factors are more localized than distal ones. They are essentially developmental influences that are relevant to a particular individual or family. In a sense, they are the case-by-case manifestation of the distal factors in one's culture. Although many distal features of a culture are objectively consistent for all members in the culture, subjective experiences of those distal factors can vary greatly. In a Third World country, everyone is not poor, nor are all of the poor people poor to the same degree. Everyone in a particular culture is exposed to and influenced by a faltering national economy, but the impact of the economy on individual lives varies widely. For example, in the Value of Children study, Turkish parents with low income and education valued their children for their ability to contribute financially to the family. However, as income and education increased, Turkish parents valued their children less for old-age security reasons and more for emotional ties. Although the distal economic influence was objectively consistent for all Turkish participants, their subjective experiences and subsequent responses varied greatly.

1.2.1. Biological Factors

Biological factors are perhaps the most significant contributor to the development of observable gender differences. Although the term biological is often used to refer to genes, in this case sex chromosomes, it also includes an organism's prenatal and postnatal environment. Although these developmental components intuitively appear to be fixed, considerable cultural variation actually exists. Differences in general nutrition and prenatal care can have profound effects on the prenatal experience. In the postnatal environment, biology and environment work together to yield cultural differences in behaviors, such as young children's sleep patterns, and attainment of developmental milestones, such as sitting and walking.

Observable gender differences are influenced by biological factors essentially from birth. Males tend to be more biologically vulnerable from the moment of conception (e.g., male:female ratio at conception of 140:100 versus 105:100 live births) and throughout the life span. However, males tend to be larger and have higher activity levels, higher basal metabolism, more muscle development, and higher pain thresholds than females at the time of birth. Physical sex differences in body structure and hormonal state are minimal throughout the preadolescent years, but the differences that exist are consistent with subsequent development. By adulthood, males tend to be taller, have a more massive skeleton, a higher muscle-to-fat ratio, higher blood oxygen capacity, increased body hair, and different primary and secondary sex characteristics. In part, these differences result from a longer growth period for boys and hormonal changes that occur after age 8. These changes generally relate to greater physical strength and stamina of males relative to females, although males tend to die an average of 6 or 7 years earlier than females. Note that these differences appear within cultures rather than across all cultures and at group rather than individual levels.

1.2.2. Socialization of Girls and Boys

Many attribute the emergence of sex-specific behaviors for boys and girls to differences in socialization. Surprisingly, many of these differences appear to be consistent across numerous cultures. Herbert Barry,

Margaret Bacon, and Irvin Child studied the socialization practices of more than 100 preindustrial societies and found that boys were generally raised with an emphasis on achievement, self-reliance, and independence; conversely, girls were typically socialized to be nurturant, responsible, and obedient. Note that these differences emerged between boys and girls within the same cultural group and not necessarily between cultures. Girls in some societies are raised to be more independent and self-reliant than boys in other cultures.

Studies in North American and Western countries have found that parents more consistently reinforce sex-typed behaviors and use higher levels of physical punishment with boys than with girls. However, cultures differ widely in how frequently physical punishments are used, and the differential treatment of males and females appears to dissipate with age. The infinite human variation in the biological and experiential influences that comprise each person's developmental history allows for many different interpretations of the same experience.

1.2.2.1. Cultural Differences in Role Socialization The impact of cultural differences in socialization was seen in the classic six-culture study, an intensive investigation of socialization practices in India, Kenya, Mexico, Okinawa, the Philippines, and the United States. Of central interest was the occurrence of aggressive, nurturant, responsible, helpful, and attention-seeking behaviors among 3- to 11-year-old children. Less gender differentiation was found in cultures in which both boys and girls helped with sibling care and household chores (i.e., Kenya, the Philippines, and the United States). In contrast, more gender discrepancies emerged in cultures in which females participated more than males in sibling care and household chores (i.e., India, Mexico, and Okinawa). Overall, girls spent more of their time in responsible, productive work, and boys spent relatively more time playing.

Further evidence for discrepant patterns of caretaking roles of males and females comes from an examination of the care-taking practices of 186 nonindustrialized societies. Thomas Weisner and Ronald Gallimore found that mothers, female adult relatives, and female children were the primary caretakers of infants. However, as infants reached early childhood, responsibility for child care fell upon siblings of both sex groups. Because young children in some societies spend as much as 70% of each day in the care of their siblings, sibling caretakers constitute an integral part of the gender role socialization process. Although mothers in such societies spend considerable proportions of their time in nonmothering, family provider roles, children in virtually every culture view mothers as responsible for children.

Fathers often fulfill the role of primary financial provider for the family, and this may limit their contributions to caretaking. Father–child interactions in many cultures are structured around play-oriented activities. Although the father's role may not appear to be essential, research has shown that fathers play an important role in the socialization of children. When fathers are absent for lengthy periods of time, boys may display effeminate behaviors, high levels of dependence, excessive fantasy aggression, and overly masculine behaviors. However, the criteria for judging behaviors as "effeminate" may vary across cultural groups. Indeed, fathers tend to exhibit higher levels of gender role socialization than mothers and they pay less attention to daughters than to sons. These findings suggest that fathers' influence is more critical for sons than for daughters.

In addition to parents, peers play an important socialization role for children in all cultures, and this role tends to increase with age. Eleanor Maccoby identified three gender-linked phenomena in children's social development with peers: gender segregation, dissimilar interaction styles, and group asymmetry (e.g., boys groups are more cohesive, sexist, exclusionary, and separate from adult culture than girls' groups). The six-culture study found that as early as age 3, children tend to seek out same-sex peers and to avoid other-sex children, and this tendency strengthens throughout grade school. In their segregated groups boys strive for dominance, take risks, and are reluctant to reveal weaknesses to each other. In contrast, girls self-disclose more, try to maintain positive social relationships, and avoid conflict. Compared with girls' groups, boys' groups are more cohesive, more exclusionary, and more separate from adult culture. Segregated groups may lead to different activities and toy choices, which may in turn lead to differences in intellectual and emotional development.

1.2.2.2. Education Across cultures, educational settings are common sources of differential gender role socialization. For example, observations of fifth-grade students in Japan and the United States revealed that teachers devoted more than half of their attention to boys, but that much of this attention was negative. Research studies have consistently found teacher expectations for student achievement to play a critical role in

predicting discrepancies in classroom performance. As one would expect, parental expectations also play an important role in academic achievement. Robert Serpell confirmed this notion with his observation in Zambia, where education is intended for boys rather than girls. Harold Stevenson and colleagues found that mothers in China, Japan, and the United States expected their sons to outperform their daughters in mathematics and their daughters' reading abilities to exceed those of their sons; however, boys and girls perform equally well in some aspects of both disciplines.

Interestingly, during the second half of the 20th century women's level of education has increased and even surpassed that of men in the United States and in several other Western countries. This is a profound shift considering that in some countries women are not permitted to participate in formal education.

2. BEHAVIORAL DIFFERENCES AMONG BOYS AND GIRLS

The aforementioned developmental considerations include some of the potential factors that contribute to different gender role behaviors. A more in-depth discussion of behavioral domains in which males and females show divergent patterns is presented here.

2.1. General Patterns of Activity

Cross-cultural consistency can be found through an examination of the characteristics of male and female group behaviors. Daniel Freeman studied 5- to 7-year-olds in eight different cultures: Australian Aboriginal, Balinese, Ceylonese, Japanese, Kikuyu, Navajo, Punjabi, and Taiwanese. He found that, relative to males, young females remained closer to home and engaged in repetitive, conversation-oriented behaviors. Males, on the other hand, frequently played in larger groups, participated in more physical activities, covered more physical space in play, and performed more unpredictable behaviors. In general, boys tend to spend most of their time with other boys, but girls often choose to interact with adults.

Children's drawings also shed light on their segregated gender preferences and gender role-consistent interests. Boys tend to produce more drawings of other boys, vehicles, monsters, and violent themes, whereas girls' drawings more often depict other girls and flowers.

2.2. General Behavioral Styles

Cross-cultural studies have generally found preadolescent males to exhibit higher rates of competitive, aggressive, and dominance-seeking behaviors and females to demonstrate higher levels of nurturance. In particular, the six-culture study found sex differences in aggression and dominance, consistent levels of aggression across many age groups, and increased physicality among the eldest boys. Indeed, fathers may play a particularly important role in the socialization of aggressive behaviors in boys. Sex differences have also emerged from a consideration of the manner in which both boys and girls engage in aggressive interactions. Initially, males display more restraint than females but are more violent when they do act. Female aggression is often marked by emotion-laden, verbal tirades.

With regard to nurturant behaviors, because infants elicit more nurturant behavior than do older children, girls who spent more time with infants displayed more nurturance than boys who did not engage in as much interaction with infants. Herbert Barry and colleagues examined socialization practices in more than 100 societies and found that generally boys are reared to achieve and to be self-reliant and independent, whereas girls are reared to be nurturant, responsible, and obedient.

2.3. Self-Esteem

Girls have often appeared to be less satisfied with being girls, and boys generally perceive themselves to be more competent than girls. However, some studies suggest that these differences tend to be domain specific and can even be reversed depending on what aspect of self-esteem is examined. David Watkins found that girls in Australia, Nepal, and the Philippines reported lower self-esteem than boys in the areas of physical abilities and mathematics. Additionally, Nigerian boys felt they were more intelligent overall than girls. However, Australian and Nigerian girls perceived themselves to be more competent than boys in reading. As one can see, male and female differences in self-esteem across cultures are not as robust as general behavioral patterns. Perhaps the inconsistencies in self-esteem across genders are due to cultural differences in the appropriateness of various behaviors. Cultures that consider certain behaviors to be more negative than do other cultures may associate them with greater maladjustment, such as low self-esteem.

3. DEVELOPMENT OF GENDER ROLES AND STEREOTYPES

Research has consistently shown that gender-specific knowledge emerges early in a child's development. In the United States, children as young as age 2 begin to stereotype objects as masculine or feminine, and by age 3 or 4 they reliably use stereotypic labels with toys, activities, and occupations. Similar patterns can be found in countries as distinct as West Africa, where girls tend to play with dolls and boys often build vehicles and weapons. Although the contextual details of children's play differ across cultures, most general play patterns are not culture specific.

3.1. Sex Trait Stereotypes

Research in the United States revealed that children's understanding of sex trait stereotypes may emerge more slowly than their stereotypic knowledge of toys and occupations. John Williams and Deborah Best developed the Sex Stereotype Measure (SSM) to assess children's understanding of adult notions of sex stereotypes. In this task, children were shown a silhouette drawing of a man and a woman while they were read stories consisting of various sex stereotype traits (e.g., aggressive, emotional, and gentle). Children were then asked to tell which silhouette each story was about. European American children's knowledge of sex trait stereotypes showed a consistent pattern of growth from kindergarten through high school. Scores showed the most dramatic gains during elementary school years and often appeared to plateau after junior high school. African American children's knowledge of the same stereotype traits also increased with age, but at a slower rate. This observation suggests that different nationalities within one country may vary somewhat in their stereotype knowledge base.

3.2. Findings across Cultures

Williams and Best used the SSM with 5-, 8-, and 11-year-old children from 25 different nations. In general, they found that the frequency of stereotyped responses increased from approximately 60% at age 5 to 70% at age 8. At all ages, items consistently associated with males were strong, aggressive, cruel, coarse, and adventurous; weak, appreciative, softhearted, gentle, and meek were consistently associated with females.

Scores were exceptionally high in Pakistan and relatively high in England and New Zealand, indicating that children in these countries had a considerable understanding of sex stereotypes. Conversely, scores were particularly low in Brazil, France, Germany, and Taiwan, suggesting consistent stereotype trait knowledge may develop more slowly in these countries. The general pattern across all countries was characterized by the onset of stereotype learning prior to age 5, accelerated learning during the school years, and functionally complete stereotype knowledge during adolescence.

Boys and girls acquired stereotype knowledge at the same rates, but there was a tendency for both groups to learn about the male stereotype at a faster rate. However, this tendency was reversed in Latin/Catholic cultures (i.e., Brazil, Chile, Portugal, and Venezuela), in which the adult female stereotype was more favorable than that of the adult male. Male stereotype items were better known than female items by both sexes in 17 of the 24 countries studied. Germany was the only country in which there was a clear pattern of greater knowledge of female stereotype items. Like the Latin/Catholic cultures, German female stereotype knowledge at age 5 was greater than male stereotype knowledge. Differences emerged at 8 years due to the persistence of greater female stereotype knowledge among German children but a change to greater male stereotype knowledge among 8-year-olds from Latin/Catholic cultures.

Five-year-olds in predominantly Muslim countries differentiated more between the sexes than 5-year-olds in non-Muslim countries. Furthermore, children in Muslim countries learned the stereotypes, particularly the male stereotype, more quickly than children living in predominantly Christian countries. Because adult stereotypes tend to be less well differentiated (i.e., overlap more) in countries in which a majority of people practice Christianity, children in these countries may acquire stereotype knowledge at a slower rate than children living in cultures distinguished by rigid boundaries that separate the characteristics generally considered to be typical of males or females.

Studying older children, Margaret Intons-Peterson found that 11- to 18-year-olds in Sweden had more similar male and female stereotypes than children in the United States. However, little overlap was found in the Swedish children's ideal occupational choices. Swedish boys were interested in business careers, whereas Swedish girls reported a preference for occupations in the service industry, such as nanny, hospital worker, and flight attendant.

4. THEORIES OF GENDER-RELATED LEARNING

Although most theories of gender role learning were developed in the United States, they consistently emphasize the gender-relevant knowledge that is available within a particular cultural context. The theories differ on many points, but they each accommodate cross-cultural patterns of gender role development.

4.1. Social Learning

From the perspective of social learning theories, gender role development is conceptualized as the accumulation of experience with socialization agents in one's environment. Socialization agents such as parents, teachers, and peers concurrently shape a child's gender-related behaviors by rewarding gender-appropriate behaviors, punishing gender-inappropriate behaviors, modeling gender-appropriate activities, selecting gender-appropriate toys, and generally treating boys and girls in a differentiated, sex role-consistent manner. Parents in the United States, and fathers in particular, have been found to treat their children differently depending on the sex of the child. For example, boys are often given more freedom and independence than girls, and they tend to receive more physical stimulation. However, studies in other countries have shown less differential treatment on the basis of sex relative to that found in the United States.

Deborah Best and colleagues observed parents and their preschool children in public parks and playgrounds in France, Germany, and Italy, and they found variation in parent–child interaction patterns across both gender and culture. In comparison to French and Italian parent–child pairs, German parents interacted less frequently with their children. Furthermore, German children showed and shared more objects with their mothers, whereas French and Italian children engaged in the same behaviors more frequently with their fathers. These observations suggest that German children, relative to those in other countries, may engage in a disproportionately greater amount of interactions with their mothers relative to their fathers, possibly fostering more rapid development of the female stereotype relative to the male stereotype. Similarly, these data suggest the opposite pattern for French and Italian children. This pattern is consistent with the aforementioned findings from the cross-cultural work of Williams and Best, which indicated that Germany was the only nation in which the female stereotype was learned more rapidly than the male stereotype and persisted at higher levels into later childhood.

It is clear that social factors play an important role in children's learning of gender roles, and there is wide variation in how girls and boys are treated across cultural groups. Surprisingly, differential treatment of boys and girls is not consistently related to gender differences in the behavioral patterns observed across cultures.

4.2. Cognitive Development

Cognitive developmental theory focuses on the way cognitive structures affect the influence of the environmental factors that are often studied by social learning theorists. Determining the role that cognitive structures play in the development of gender knowledge and gender-typed behaviors is a difficult task due to the changing nature of children's cognitive structures across childhood. As these structures change with age, so too does the manner in which children interpret and remember gender-relevant experiences.

Consistent with Jean Piaget's stage theory of cognitive development, Ronald Slaby and Karin Frey identified four developmental stages that characterized children's gender knowledge acquisition in North America. Prior to age 2, children did not distinguish between men and women, but in the second stage by age 2 or 3 they used physical characteristics to differentiate between the sexes. By the third stage at approximately 4 years of age, children understood that gender was stable, and by age 5 or 6 they knew that gender remained constant.

Ruth and Robert Munroe and Harold Shimmin tested the predictions of cognitive developmental theory in a study of children in societies marked by differing degrees of gender differentiation. They predicted that in societies with high gender differentiation (i.e., Kenya and Nepal), children would progress through the cognitive gender stages at a faster rate than children living in less differentiated settings (i.e., Belize and Samoa). Contrary to these predictions, attainment of gender classification capabilities was not related to cultural variations in gender differentiation. Although the degree of gender differentiation varied between these cultures, the discrepancies within each culture may have been sufficient to promote the development of gender discrimination skills. Furthermore, it is possible that the extensive gender differentiation in Kenya and Nepal may have been superfluous for children's learning.

5. GENDER AT THE ADULT LEVEL

It is apparent from the research on the development of gender roles and stereotypes that children's cultural environment influences much of what they learn. Children are socialized by adults whose social behaviors, gender stereotypes, and self-perceptions reflect differences between men and women.

5.1. Sex Role Ideology

Sex role ideology concerns an individual's beliefs about men's and women's proper role relationships, ranging along a continuum from traditional to modern. The traditional end of the continuum is characterized by beliefs that men are more "important" than women and that women should be subservient to men. Modern ideologies are more egalitarian and are consistent with the notion that women and men are equally important and that neither sex has the right to dominate the other.

Because of the juxtaposition of traditional and modern ideologies among its inhabitants, India has been a fascinating nation in which to study adult sex roles. Khavita Agarwal, David Lester, and Nisha Dhawan asked university students in India and the United States to identify the qualities men and women should possess in their respective cultures. They found that North American students expressed more modern views than Indian students; however, across both cultures, women were more liberal (modern) than men. Rehana Ghadially conducted a more in-depth analysis of Indian women and found that those with the most modern sex role ideologies grew up in nuclear families, had educated mothers, and worked in professional disciplines.

Atsuko Suzuki conducted similar studies in Japan and the United States and found education and job status to be strong predictors of sex role attitudes among women. North American women with jobs held more egalitarian views than women without jobs. In Japan, women with career-oriented professional jobs reported more egalitarian views than all other women.

Judy Gibbons, Deborah Stiles, and Gina Shkodriani studied attitudes toward family roles and gender among adolescents studying in The Netherlands. The students originated from 46 different countries and formed two groups based on origin: individualistic, wealthier countries and collectivistic, less wealthy countries. Students from this second group of countries reported more traditional attitudes than students from the first group, and across both groups, boys reported more traditional attitudes than girls.

In studies of sex role ideology, the United States has often served as a comparison group, and the beliefs of U.S. adults are generally found to be more liberal than those of adults of other nations. However, this was not the case in a 14-country study of the attitudes of university students by Williams and Best. Northern European countries (i.e., The Netherlands, Germany, Finland, and England) had the most modern ideologies, the most traditional attitudes were found in African and Asian countries (i.e., Nigeria, Pakistan, India, Japan, and Malaysia), and the United States was in the middle of the distribution. With the exception of Malaysia and Pakistan, women reported more modern attitudes than men. However, within a given country, men's and women's sex role ideology scores were highly similar; that is, greater differences existed between cultural groups (men and women combined) than within cultural groups (men compared to women).

In cross-cultural research it is necessary to show that differences between cultural groups are related to other cultural variables before one can conclude that the differences of interest are due to cultural factors. In the 14-country study noted previously, sex role ideology scores were related to socioeconomic development, supporting the role of cultural factors. Specifically, sex role ideology scores were more modern in more developed nations, in countries in which Christianity was more predominant, in more urbanized countries, and in countries at higher latitudes (i.e., countries at higher latitude tend to have greater socioeconomic development).

5.2. Gender Stereotypes

Gender stereotypes are the psychological traits believed to be more characteristic of one sex relative to the other. These stereotypes are closely related to sex role ideologies, and they have often been used to account for variations in sex role ideologies across cultures. Williams and Best presented university students in 27 countries with the Adjective Checklist, a standard personality measure consisting of 300 person-descriptive adjectives. Students indicated whether, in their culture, each adjective was more characteristic of men, more characteristic of women, or not differentially associated with either sex. Students generally agreed about the characteristics differentially associated with men and women, but differences between the male and female

stereotypes did emerge. Male and female stereotypes varied the most in The Netherlands, Finland, Norway, and Germany and least in Bolivia, Scotland, and Venezuela. Additionally, greater stereotype variation was found in Protestant compared to Catholic countries, in more developed countries, and in countries higher in individualism.

Analysis of the qualitative aspects of the male and female stereotypes revealed that the male stereotype items were stronger and more active across cultures. Across all countries, items representing dominance, autonomy, aggression, exhibition, and achievement were associated with men, and nurturance, succorance, deference, and abasement were consistently chosen to describe women.

Recent work examined these stereotype findings in terms of the Five Factor Model of personality, an approach based on the assumption that five dimensions can fully capture the aspects of personality. The pancultural male stereotype, which consisted of items consistently associated with men across most cultures, was found to be higher on the dimensions of extraversion, conscientiousness, emotional stability, and openness, whereas the female stereotype was higher on the dimension of agreeableness.

In contrast to these cross-cultural generalities, variation between the countries emerged as well. No consistent pattern emerged for stereotype favorability, with the male stereotype items being more favorable in some countries (e.g., Japan, South America, and Nigeria) and the female stereotype items exceeding the male items in others (e.g., Italy, Peru, and Australia). Greater variation in the perceived strength and activity of the male and female stereotypes was found in countries that were less socioeconomically developed and had lower literacy rates and lower percentages of women attending universities. These findings suggested that economic and educational advancements were accompanied by reductions in the tendency to view men as stronger and more active than women. Note, however, that these effects were only reduced, not eliminated.

5.3. Masculinity/Femininity of Self-Concepts

The paired concepts of masculinity and femininity essentially mean manlike or woman-like, and a person can be characterized as masculine or feminine along a number of dimensions, such as dress, mannerisms, or tone of voice. One aspect of masculinity and femininity relates to self-concepts and the degree to which individuals incorporate traits that are differentially associated by men or women into their self-concepts. Researchers have used a variety of measurement techniques to assess individual differences in the masculinity and femininity of self-concepts. For example, some researchers have used self-descriptive questionnaires, some have analyzed only socially desirable characteristics, and others have examined gender-associated characteristics without considering social desirability. Certainly, a conjunction of the three would provide a more comprehensive examination.

Valid measurement is a continuing concern for researchers studying the similarities and differences between various cultural groups. Measurement problems can arise when a scale is developed in one culture and then applied in other cultures without attention to the validity of the instrument in the new context. A prime example of questionable measurement validity occurred in a cross-cultural study of masculinity and femininity by Robert Helmreich and Janet Spence. They used the Personal Attributes Questionnaire, which contains traits associated with either males or females in the United States, to compare the self-descriptions of men from Brazil and the United States. They found that Brazilian men endorsed more feminine than masculine traits, whereas men from the United States endorsed the opposite pattern. On the surface, these results suggested that men in the United States were more masculine than men in Brazil, and, surprisingly, that Brazilian men may have been more feminine than masculine. However, these accounts do not consider cultural differences in the definitions of masculinity and femininity. It is equally plausible that, from the Brazilian perspective, the self-descriptions of men in the United States are more feminine than those of Brazilian men.

The aforementioned difficulties were addressed in a study by Williams and Best in which the masculinity and femininity of items were determined within each culture. These authors assessed the masculinity and femininity of the self-concepts of university students in 14 countries. Participants selected the items from the Adjective Checklist that they believed described their actual and ideal self-concepts (i.e., who they really were vs who they would like to be). These descriptions were scored according to university student ratings of the masculinity and femininity of each item in prior studies in each of the 14 countries. Therefore, the stereotyped items reflected culture-specific conceptualizations of

masculinity and femininity. As one would expect, the self-concepts of men were more masculine than those of women in each culture. Surprisingly, members of both sexes rated their ideal selves as more masculine than their actual selves. Considering the preceding discussion of the affective features of male and female stereotypes, this trend of desired masculinity may reflect a universal desire for a stronger, more active self.

Variation did exist in the self-concepts of men and women across the cultures studied, but this variation was not related to meaningful cultural comparison variables, such as socioeconomic development. When self-concepts were scored in terms of the relevant culture's standards for masculinity/femininity, no evidence was found to suggest cross-cultural differences in the degree of masculinity among men or femininity among women.

In contrast, the discrepancy between ideal and actual self-concepts varied considerably across cultures, and much of this variation was systematic. In particular, the differences between men's and women's actual and ideal self-concepts were smaller in cultures characterized by greater socioeconomic development, higher rates of female employment outside the home, higher rates of female enrollment in universities, and modern sex role ideologies.

6. RELATIONS BETWEEN WOMEN AND MEN

6.1. Mate Preferences

Social scientists have often considered mate preferences to be culture bound and therefore inconsistent across cultures. However, findings by David Buss and associates suggest quite the contrary. In the most extensive investigation of mate preferences to date, these researchers collected responses from more than 10,000 individuals in 33 countries. Using two lists of potential mate characteristics, participants indicated their preferences by ranking or rating the items. Surprisingly, a high degree of similarity in the preferred characteristics of mates emerged for women and men across cultural groups. For the 13 items considered, both sexes ranked "kind and understanding" first, "intelligent" second, "exciting personality" third, "healthy" fourth, and "religious" last. However, women tended to place a higher value on a potential mate's earning capacity than did men, and men generally valued physical appearance more than women. In a reanalysis of the Buss data, Alice Eagly and Wendy Wood also found that men valued a good cook and housekeeper more than women.

In contrast to the aforementioned similarities, cultural differences were found for nearly every item. The greatest variation was found for the item "chastity." Groups in China, India, and Iran placed considerable importance on this item, but those in northern Europe viewed it as relatively unimportant. In general, men valued chastity more than did women.

Considering the remarkable degree of similarity across samples, Buss suggested that substantial unity exists in human mate preferences, and that this unity may be regarded as "species typical." However, no sample was identical to another, and each group showed unique tendencies in the ordering of characteristics, indicating at least modest degrees of cultural variation.

Not only are there similarities along with some differences in preferences for human mate characteristics, but the recent International Sexuality Description Project of 16,288 people across 10 major world regions revealed similar findings regarding desire for sexual variety. Universal sex differences were found, with men possessing greater desire than women for a variety in sexual partners, for less time to elapse in a relationship before consenting to sexual intercourse, and for more active seeking of short-term sexual partners. However, for each of these sex differences, there was significant variation by world region. For example, compared with other regions of the world, East Asian men and women were lower in their desire for more than one partner in the next month.

6.2. Romantic Love

Romantic love and intimacy, as with mate preferences, are assumed to be primarily under cultural influences. In general, romantic love is valued in more modern cultures, such as the United States, in which individualism is emphasized and few have extended family ties. However, romantic love is viewed as less essential in cultures in which strong family ties reinforce the family relationships brought about through marriage, such as in Japan. There does appear to be some degree of universality when one considers an analysis of the ethnographies and folklore materials from 166 societies. This review found at least one incidence of passionate love in 147 (88%) of the cultures examined, suggesting that the potential for romantic love may be a human universal that is overwhelmed by other cultural variables.

A study by Josephine Naidoo addressed these issues with Asian Indian immigrants in Canada by examining

generational changes in the attitudes reported toward love and marriage. Although nearly two-thirds of the immigrants had arranged marriages, a substantial proportion of them believed that "love marriages" were an option for their offspring. More than 70% of the second generation believed love should precede marriage and wanted their children to have greater freedom in mate selection than they had experienced.

6.3. Harassment and Rape

Kaisa Kauppinen-Toropainen and James Gruber conducted one of the few cross-cultural studies of male harassment and hostility toward women. After examining blue-collar and professional women in Scandinavia, the former Soviet Union, and the United States, these authors found that Americans reported the most woman-unfriendly experiences. Relative to Americans, Scandinavians reported fewer job-related or psychological problems, more autonomy, and better work environments. Professional women in the former Soviet Union reported fewer negative experiences than blue-collar women but more than professional women in the United States and Scandinavia.

Colleen Ward and colleagues conducted a comprehensive assessment of attitudes toward rape victims with university students from 15 countries. Cross-cultural variation in attitudes toward rape victims followed patterns of attitudes toward women more generally. That is, countries with more modern sex role ideologies reported more favorable attitudes toward victims, whereas less favorable attitudes emerged in countries with lower percentages of women in the labor force and lower literacy rates. In particular, more favorable attitudes were reported by university students in Germany, New Zealand, and the United States, and less favorable attitudes were found in India, Malaysia, Mexico, Turkey, and Zimbabwe.

7. GENDER DIFFERENCES AT THE CULTURAL LEVEL

7.1. Masculinity and Work-Related Values

Using attitude survey data from thousands of IBM employees who were predominantly male, Geert Hofstede compared work-related values in 40 countries. One scale derived by Hofstede focused on the extent to which values of assertiveness, money, and things prevail instead of values of nurturance, quality of life, and people. This scale was labeled Masculinity (MAS) due to the fact that male employees assigned greater importance to the first set of values, whereas females assigned greater importance to the second set. A MAS index was computed for each of the countries in Hofstede's study; the five highest rated countries on the MAS dimension were Japan, Austria, Venezuela, Italy, and Switzerland, whereas the five lowest rated were Sweden, Norway, The Netherlands, Denmark, and Finland. High MAS countries were characterized by stronger beliefs in independent decision making, stronger achievement motivation, increased job stress, and a view of work as central in people's lives.

The significance of the MAS dimension is unquestionable; however, many researchers have challenged the appropriateness of labeling the scale Masculinity. This label leads to the expectation that variation on the MAS scale would be associated with cross-cultural variation on other gender-related constructs. Contrary to these expectations, Williams and Best have not found relationships between their sex stereotype or masculinity/femininity measures and Hofstede's MAS dimension. Other researchers have also failed to find the expected relationships.

7.2. Gender Division of Labor

Conceptions of masculinity and femininity vary widely across cultures, but two universals are plausible: (i) To varying degrees, every society assigns traits or tasks on the basis of sex, and (ii) the status of women is inferior to the status of men in every society. As one would expect based on these generalizations, extensive differences do exist in the work roles of men and women. Examining jobs and tasks in 244 societies, Roy D'Andrade found that men were involved in hunting, metal work, and weapon making and tended to travel further from home than women. Women were responsible for food preparation, carrying water, caring for clothing, and various child-rearing responsibilities. Although women's subsistence activities generally included child-rearing demands, some did hunt in societies in which this activity did not compete with child care. The strong sex segregation for child-rearing duties was mirrored by another study that found that men were significant caretakers in only 10% of the 80 cultures examined. However, both sexes seemed to be flexible enough to adapt to a range of socioeconomic roles.

Today, women account for a substantial proportion of the world's labor force. With decreases in infant mortality and fertility, women now spend less time in child-rearing

roles. Furthermore, technological advances have allowed women in many parts of the world to separate childbearing from child-rearing and thereby contribute to the family through jobs outside the home.

However, women's increased autonomy has not been paralleled by increased acceptance or equality. For example, in a 56-country study of labor trends from 1960 to 1980, the job market was marked by declines in women's occupational opportunities and increases in sex segregation. When measured by per capita gross national product and women's level of education, modernization was associated with increased segregation of the sexes. Additionally, increased workplace involvement for women correlated with decreases in total fertility rate. Women continue to be disadvantaged in the workplace, most overtly through persistent salary discrepancies that favor men. In addition to women's lower salaries is evidence suggesting that women prefer traditionally female jobs, especially those offering extensive contact with other people. Moreover, these jobs tend to be low paying. On the contrary, men tend to prefer jobs with high income and promotion opportunities.

Even in the countries with the highest proportion of females in the labor force, women continue to face inequality within the home. Studies in several North American and European countries have found that women perform a majority of the housework, regardless of the extent of their occupational demands. Along with children and larger homes comes reduced male involvement in domestic chores. This is surprising in light of the previous suggestion that increases in education and income are associated with more modern sex role views (i.e., equality in the workforce). However, studies suggest that systematic differences in sex role ideology persist in these more modern countries. For example, in the United States, Great Britain, West Germany, and Austria, people with relatively high levels of education and women with employed husbands indicated less support for efforts to reduce gender inequality compared to those with less education and women without employed husbands. Such findings suggest that even the subjective change in perceived life quality associated with improved socioeconomic conditions may be greater for men than women.

8. CONCLUSION

Sex differences continue to pose challenging questions for social and biological scientists, and with increased interest in the role of culture, new questions will certainly emerge. Although males and females are biologically more similar than different, they lead different lives, regardless of their society's tendencies toward traditional or egalitarian views.

Considering the degree of gender stereotype similarity across cultures suggests a pancultural model in which cultural nuances produce relatively small variations around general tendencies. Within this model, the relatively modest biological differences are perceived to set the stage (e.g., women bear children, men have greater physical strength) and subsequently lead to a division of labor such that women are responsible for child-rearing and other domestic duties, whereas men are providers (e.g., hunting) and protectors. From the perspective of this model, gender stereotypes evolved to support this division of labor. Gender stereotypes lead to an implicit assumption that members of either sex have or can develop characteristics that are consistent with one's assigned role. Once they are established, such stereotypes can guide differential sex role socialization for boys and girls and lead boys to become more independent and adventurous, whereas girls become more nurturant and affiliative. Additionally, such boundaries heighten the salience of people who differ from societal expectations of gender-appropriate behavior, possibly creating social difficulties for those people. Although the accuracy of such a model is difficult to test due to its reliance on evolutionary processes, its usefulness is apparent because it demonstrates a means by which people in widely different cultures can come to associate one set of characteristics with men and another with women, with only minor deviations from these central themes. The relatively modest biological differences between the sexes appear to be enhanced or diminished by cultural practices; however, even small differences in gender roles and behaviors can have important cultural implications.

See Also the Following Articles

Gender and Education ■ Gender Role Development ■ Occupation and Gender ■ Sexual Harassment

Further Reading

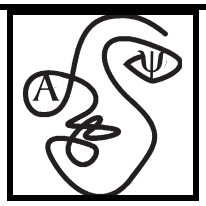
Best, D. L., & Thomas, J. J. (2003). Cultural diversity and cross-cultural perspectives. In A. H. Eagly, A. E. Beall, & R. J. Sternberg (Eds.), *The psychology of gender* (2nd ed.). New York: Guilford.

Edwards, C. P. (1992). Cross-cultural perspectives on family–peer relations. In R. D. Parke, & G. W. Ladd (Eds.), *Family–peer relationships: Modes of linkages*. Hillsdale, NJ: Erlbaum.

Hoyenga, K. B., & Hoyenga, K. T. (1993). *Gender-related differences: Origins and outcomes*. Boston: Allyn & Bacon.

Maccoby, E. E. (1998). *The two sexes: Growing up apart, coming together*. Cambridge, MA: Harvard University Press.

Yoder, J. D. (2003). *Women and gender: Transforming psychology* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.



Gender and Education

Linda M. Raffaele Mendez
University of South Florida, Tampa, Florida, USA

1. What Is Gender Equity?
2. Gender Equity in Schools: Issues of Concern
3. Increasing Gender Equity: Strategies for Educators
4. Summary
   Further Reading

GLOSSARY

academic retention A school's decision to keep a student in a grade for another year rather than having the student matriculate to the next higher grade.

the boy code A set of unwritten rules that require boys to act tough, independent, and confident even in situations in which they feel frightened or unsure.

classroom dynamics The interactions occurring between individuals in a classroom.

disorders of body image Negative perceptions and attitudes about one's own body (often manifest as being too fat among girls or not having enough muscle among boys).

externalizing behavior Disruptive behavior, such as noncompliance or aggression.

gender equity Girls and boys experiencing equal opportunities in educational settings.

gender stereotypes Commonly held beliefs about characteristics deemed appropriate for males and females.

relational aggression A form of aggression that damages others' relationships (e.g., social exclusion and rumor spreading).

self-contained classroom A classroom serving only students in special education, most commonly because of a perception that the students' needs cannot be accommodated in the general education classroom.

unified discipline approach A method of handling discipline that involves partnerships and consistency between schools, families, and communities rather than conflicting or contradictory policies between these groups.

In most industrialized countries, laws have been established to mandate equal educational opportunities for boys and girls. For example, in 1972, the United States passed Title IX, a federal law that mandates that no person be discriminated against on the basis of sex in any educational program receiving federal funds. One might assume, then, that gender does not have much of an impact on education (i.e., that there are no appreciable differences in how girls and boys experience school). In contrast to this assumption, research shows that gender can have a major impact on a child's educational experiences. This article reviews the literature on gender and education with a particular focus on the issue of gender equity.

1. WHAT IS GENDER EQUITY?

When applied to education, the term gender equity refers to girls and boys experiencing equal opportunities in educational settings. The term not only refers to academic learning but also applies to a much broader range of activities and experiences that are a part of the educational process. In an equitable educational environment, boys and girls have similar opportunities to

©2004 Elsevier Inc.
All rights reserved.

master the academic curriculum, experience a sense of comfort and belonging, have interactions with teachers and peers that promote both competence and confidence, be recognized for their contributions, and participate in extracurricular activities that are enjoyable and build self-esteem. In order to create such an environment, educators must understand and appreciate boys' and girls' unique developmental issues and needs within schools. Educators must also seek to eradicate gender stereotypes that limit opportunities for youth (e.g., that girls cannot be successful in math or that boys should not be interested in cooking or sewing). Funding (e.g., of athletic teams) and access (e.g., to vocational programs that lead to higher paying jobs) also must be given careful consideration in a gender-equitable educational environment.

2. GENDER EQUITY IN SCHOOLS: ISSUES OF CONCERN

2.1. Concerns for Girls

Much of what has been written about gender equity in education has focused attention on educational inequity for girls. Frequently noted concerns include girls receiving less teacher attention than boys, biased curricular materials presenting women in subordinate roles to men and/or failing to recognize the significant accomplishments of women (e.g., excluding the work of female scientists from textbooks), attitudes toward sexual harassment that accept such behavior as being a normal part of the adolescent school experience, and stereotypes indicating that girls are less able than boys in mathematics and the sciences.

Several publications from the early to mid-1990s did much to raise awareness of gender inequity in schools. For example, David and Myra Sadker, in their book, *Failing at Fairness: How Our Schools Cheat Girls*, summarized research showing that boys stand out in the typical classroom (receiving both more positive and more negative attention than girls), whereas girls quietly do their work and blend into the classroom landscape. It is notable that research has shown that low-achieving girls are less likely to be referred for special education services than low-achieving boys, even when their levels of achievement are similar. This is because low-achieving boys tend to manifest more behavior problems, which makes them stand out more. Also, when girls are in need of a self-contained classroom experience, especially a classroom for children with emotional or behavioral problems affecting their learning, there tend to be very few appropriate classroom placements since these classes are dominated by boys.

Importantly, many of the types of problems that are more common among girls than boys (particularly in adolescence) and may be related to diminished opportunities for classroom learning (e.g., adolescent depression, eating disorders, disorders of body image, sexual harassment, and relational aggression) often do not come to the attention of teachers or other school personnel because they tend not to disrupt classroom activities. Thus, teachers may not know that girls in their classrooms are underperforming because of unaddressed social or emotional issues. Boys are much more likely to act out their problems than girls and thus are also more likely to come to the attention of those who might be able to provide assistance (e.g., school psychologists and counselors).

2.2. Concerns for Boys

Systematic inquiry into the ways in which boys are underserved in schools has only begun to emerge in the past 5–10 years. One important source of concern for boys centers on the fact that most educational environments, while becoming increasingly tolerant of girls engaging in traditionally male activities (e.g., playing physical contact team sports), are still highly stereotyped for boys. For example, William Pollock, in his book, *Real Boys: Rescuing Our Sons from the Myths of Boyhood*, noted that boys are more likely than girls to be criticized by teachers for expressing feelings of fear (e.g., on the first day of school) or affection (e.g., hugging another child). Male peers also are likely to chastise boys for any behavior that is perceived as feminine or as being "like a girl" (e.g., boys whose athletic skills are not well developed might be criticized as "throwing like a girl"). These observations exemplify the fact that boys are still restricted by what Pollock calls "the boy code," which stipulates that boys must act tough, independent, and confident even in situations in which they feel frightened or unsure. Most educators are much more likely to tolerate cross-sex interests and behavior from girls than boys.

Another important concern regarding educational equity for boys is related to the observation that so many more boys than girls experience academic or behavioral failure in school. Boys are considerably more likely than girls to be retained, to be identified as having special needs and placed in a special education program, and to

experience school discipline (e.g., suspension and expulsion). Whether such gender differences are related to biological or social factors (or a combination of the two) is unclear. Nonetheless, given these outcomes, the possibility must be considered that the types of educational environments currently existing in our schools are not meeting boys' unique needs. Some researchers have noted that many boys begin kindergarten behind their female peers in terms of fine motor skills, verbal skills, attentional capabilities, and other important indicators of readiness for formal learning. This appears to be due to slower overall maturation in boys compared to girls. However, instead of conceptualizing boys as having different needs in early elementary school than girls (e.g., more opportunities for learning through gross motor as opposed to fine motor activities), educators tend to be quick to label boys as deficient, leading to early elementary retention in grade or referral for special education testing. If educational environments gave more careful consideration to boys' unique needs, gender differences in educational failure might be reduced.

3. INCREASING GENDER EQUITY: STRATEGIES FOR EDUCATORS

There is much that educators can do to ensure that children's educational experiences are consistent with their needs. With regard to enhancing gender equity in educational settings, there are three primary areas to be considered: classroom dynamics/instructional style, the curriculum, and behavior management/intervention.

3.1. Classroom Dynamics/ Instructional Style

As noted previously, David and Myra Sadker's work focused on the fact that, in general, girls receive less teacher attention in the typical classroom than boys. As such, boys experience the benefits of a more intense educational climate, receiving both more attention from teachers overall and more precise attention (e.g., greater feedback when responding to a teacher-posed question) than girls. It is likely that the discrepancy in the way teachers treat boys and girls is related to gender differences in activity level and interactional style, with boys demanding more attention through their behavior than girls. Nonetheless, differences between boys and girls in teacher attention could be reduced if teachers were more aware of these issues. Thus, one of the primary means for increasing gender equity in classrooms is to educate school staff about the ways in which they may be treating boys and girls differently. Because at least some of this behavior is unintentional, it can be helpful for teachers to have someone such as the principal, school psychologist, or guidance counselor (i) educate teachers about the truths and myths of gender differences (e.g., boys are not inherently better at math and science than girls), (ii) train teachers directly in gender-fair teacher–student interactional practices (e.g., through modeling and/or role-playing gender-fair expectancies, practices, and attitudes), and (iii) observe and provide feedback on the degree to which teachers are engaged in gender-fair teaching methods and strategies. Many teachers may not realize that they never ask a girl to help if they are having a problem with the TV or VCR or that they communicate to boys that they should act tough even if they are hurt or scared. Bringing these issues to awareness is the first step in creating positive change.

3.2. The Curriculum

A second important issue to address is the curriculum. It is important that educational materials used in schools (e.g., textbooks, posters, media, and Internet resources) represent the diversity of the human experience. In terms of gender, this includes depicting both men and women of various ethnic and racial backgrounds in positions of leadership and achievement. Students should have the opportunity to learn about the important contributions that both women and men have made to society and to see depicted in their educational materials role models who help them to envision possibilities for themselves that defy gender role stereotypes. It also includes balancing the number of books or other educational materials created by men and women as well as the gender of the strong central figure in these materials.

In addition to carefully considering educational materials, there are a number of other ways for school administrators to reflect to students, staff, and the larger community that students at their school will not be limited by traditional gender role stereotypes. These include developing nontraditional career mentoring programs, providing similar levels of financial support for girls' and boys' athletic teams, and establishing awards programs that recognize teachers for excellence in promoting gender equity and students for significant accomplishments in areas traditionally considered to be more appropriate for the opposite

gender (e.g., girls in science or math and boys in writing or the arts). Additionally, at the middle and high school levels, it is particularly important for schools to find proactive ways to address the problems of teenage pregnancy, sexual harassment, and sexual violence. These are concerns that many school administrators (and parents and teachers as well) do not believe should be discussed in schools. Nonetheless, research demonstrating the disproportionate negative impact of these issues on young women's schooling experiences is cause for concern. It also is recommended that schools have a policy prohibiting sexual harassment. It is particularly important that this policy include student-to-student sexual harassment because this is the primary form of sexual harassment that students experience in school.

Finally, with regard to the curriculum, educators must address the question of how well the curriculum at each grade level is meeting the needs of both boys and girls. For example, if twice as many boys as girls are being retained in kindergarten, the question of whether the current kindergarten curriculum is meeting the needs of boys must be raised. One sex should not set the standard to which the other sex is compared. Differences in development and unique needs must be recognized. Similarly, if girls are not opting to take upper-level science and mathematics courses in high school, the reasons behind this should be investigated. Are those courses being taught in a way that optimizes girls' learning potential? Are girls choosing other alternatives because they have been socialized (even inadvertently) to believe that they cannot be successful in these courses? Do they realize that their college options may be limited without this coursework in high school? These types of questions can help educators assess the degree to which they are meeting unique gender needs at different levels of the curriculum.

3.3. Behavior Management/ Intervention

It has been noted that boys are greatly overrepresented with regard to grade retentions, referrals to special education, and school disciplinary action compared to their female peers. In part, this is because many more boys than girls demonstrate problems with externalizing behavior in the classroom (e.g., noncompliance and aggression). Thus, in order to meet boys' unique needs, it is very important, from a preventive standpoint, to put in place behavioral strategies that attempt to curtail disruptive classroom behavior. Most school suspensions—which have a disproportionate impact on boys—are for classroom disruptions and disobedience. Research has shown that such behaviors can be reduced through preventive programs that involve positive behavior management, social skills instruction, a unified discipline approach (involving partnerships between schools, families, and communities), and shared expectations for socially competent behavior.

It also is important to recognize that in schools, problems that cause disruption in the classroom are those that are most likely to be recognized and further assessed. Thus, the fact that more boys than girls express problems outwardly (e.g., through fighting) while more girls than boys tend to internalize their problems means that boys' problems come to the attention of teachers more often than girls' problems. Boys are thus more frequently referred to the school's student support team for assistance. Educators need to become more aware of the types of problems that are more common in girls than boys (e.g., learning problems that occur in the absence of behavioral problems, depression, eating disorders, and relational aggression), the warning signs associated with these problems, and the necessity of involving members of the student support team in intervening with girls in these situations.

Additionally, it is important to recognize that there are sex differences in the degree to which boys and girls feel comfortable sitting down with an adult to discuss their thoughts and feelings. William Pollock noted that boys have been socialized to believe they should handle their problems on their own rather than reach out to others for help. As a result, it is less likely that boys, compared to girls, will readily open up to a teacher about their feelings when asked about a problem. This may be particularly true when the teacher is female and/or of a different racial background. Pollock's research suggests that in such situations, boys may feel more comfortable focusing on an activity (e.g., games and role plays) and having discussions of thoughts and feelings arise out of the activity rather than simply being asked to share what is on their minds.

4. SUMMARY

This article has described how gender impacts educational experiences and what can be done to

increase gender equity in education. The resources listed in Further Reading give interested readers additional insight into how gender role stereotyping can limit opportunities for children and what can be done to ensure that both boys and girls receive equal opportunities to maximize their own personal potential.

See Also the Following Articles

Academic Interventions ■ Aggression and Culture ■ Attention Deficit Disorders: School-Based Interventions ■ Educational Achievement and Culture ■ Gender and Culture ■ Learning

Further Reading

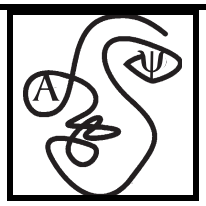
Gurian, M. (1999). *A fine young man: What parents, mentors, and educators can do to shape adolescent boys into exceptional men*. New York: Tarcher/Putnam.

Jossey-Bass. (2002). *The Jossey-Bass reader on gender in education*. San Francisco: Jossey-Bass.

Pollock, W. (1999). *Real boys: Rescuing our sons from the myths of boyhood*. New York: Owl Books.

Sadker, M., & Sadker, D. (1995). *Failing at fairness: How our schools cheat girls*. New York: Touchstone.

U.S. Department of Education. (2000). *Trends in educational equity of girls and women* (NCES Publication No. 2000-030). Washington, DC: National Center for Education Statistics. [Available online at http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2000030]



Gender, Overview

Florence L. Denmark
Pace University, New York, New York, USA

1. The Difference between Gender and Sex
2. Gender Acquisition in Childhood
3. The Harmful Effects of Gender Stereotyping
4. Gender's Sexual Component
5. Marriage and Parenthood
6. Gender and Old Age
7. Feminism
8. Suggestions for Applied Psychology
   Further Reading

GLOSSARY

ageism The discrimination of individuals based on societal stereotypes of age and aging.
blatant sexism The transparent harmful and unequal treatment of an individual based on his or her sex.
covert sexism The intentional, hidden, and hostile sexism focused on assuring the failure of a member or members of a particular sex.
cultural feminism Type of feminism that emphasizes the qualities and characteristics generally applied to women, such as the qualities of caring, nurturing, and concern for the needs of others, which society usually devalues and ignores.
feminism The belief that women and men are equal and should be equally valued as well as have equal rights.
gender A social construction based primarily on sex, consisting of traits, interests, and behaviors ascribed to each sex by society.
liberal feminism Type of feminism that focuses on the promotion of legal and social equality between men and women; advocates for a revision of values, laws, and mores associated with the inequality of women in society.
radical feminism Type of feminism that focuses on men's control over and oppression of women, which remains the primary source of women's oppression economically and socially.
sex The biological differences in genetic composition and reproductive structures and functions of men and women.
sex discrimination The unequal and harmful discrimination of individuals based on their sex.
sexism The experience of discrimination based on sex.
socialist feminism Type of feminism that focuses on social relationships and the preservation and promotion of male dominance in social institutions and that seeks to exact economic institutional change in order to achieve equality between men and women.
subtle sexism The expression of internalized negative cultural messages regarding the inferiority of members of one sex.

Gender remains an important, and often overlooked, consideration in psychological research, theory development, and practice. The cultural, structural, and cognitive components of gender contribute to the life experiences of both women and men. Cultural pressures, along with gendered cognitions, affect people of both sexes differently throughout the life cycle. Gender stereotypes and expectations regarding gender appropriate behavior bombard men and women from birth. Often these gender role expectations create difficulties for individuals socially, sexually, physically, and psychologically. The concept of feminism, loosely

©2004 Elsevier Inc.
All rights reserved.

defined as a belief in the equality of women and men, provides a framework for the treatment and study of individuals of both sexes. The articles contained in this section address a broad range of gender issues, including gender development, cultural impact on gender, gender and sexuality, homosexuality, homophobia, sexual harassment, rape, and issues pertaining to women's health in old age. Sensitivity to the cultural, structural, and cognitive origins of gender and gendered behavior contributes to the field of applied psychology in considering the different experiences of men and women throughout the life cycle.

1. THE DIFFERENCE BETWEEN GENDER AND SEX

The terms sex and gender are often misunderstood and misused. The term sex refers to the biological differences in genetic composition and reproductive structures and functions of almost all organisms. While the actual number of sexes existing in non-mammalian species remains controversial (some organisms are asexual, or have both male and female sex characteristics), scientists currently recognize two biological sexes in mammals, male and female.

Human males and females differ in their chromosomal makeup, genetically, hormonally, and neurologically. These differences create confusion and widespread misunderstanding regarding male and female biological processes and the ways in which these processes interrelate. These misunderstandings often create unfounded beliefs and confusion regarding the distinction between the male and female sex. For example, many individuals believe that females alone produce the hormones estrogen and progesterone. In addition, many hold the belief that only males produce testosterone. However, these hormones exist in both sexes to varying degrees depending on individual genetic composition, development, and biological sex.

While species may be divided in terms of sex, only humans are differentiated by both sex and gender. Gender, although based on primarily on sex, remains a social construction. Societies ascribe various gender differences based on traits, interests, and behaviors. Social context, the social sphere in which an individual exists and interacts, remains the greatest determinate of the designation of gender, that of "masculine" and "feminine." Personal history, immediate situation, society, and culture comprise the social context in which individuals behave and interact. This interaction, the reciprocal activities that encompass being male and female, continues throughout day-to-day existence and society as individuals abide by society's prescriptions for gender.

1.1. The Construction of Gender

Gender remains a socially constructed concept that refers to the creation and explanation of the differences between girls and boys and women and men. The socially constructed nature of gender debases the concept of absoluteness in the biological differences between women and men. Therefore, biological differences between the sexes account for very little of the observations of different behaviors and attitudes among women and men. Rather, the social construction of gender and the societal pressures regarding behavior and attitudes deemed gender appropriate contribute to the manifestation of gender in individuals. The acquisition of gender-appropriate behavior occurs through several psychological processes involved in the etiology of beliefs regarding the adoption and maintenance of gender and gender-related activities.

Currently, there exist three approaches to conceptualizing the evolution of the construct of gender: the cultural approach, the structural approach, and the cognitive approach. The cultural approach suggests that the acceptance of culture and society's values and norms occurs through societal rewards and that punishments contribute to the appearance of gendered behavior. Through rewards and punishments, children become socialized to behave according to cultural and societal demands. The acceptance of cultural gender values continues throughout the life cycle and leads to the reinforcement and eventual perpetuation of gender stereotyping.

Another approach, the structural approach, suggests that the constant evolution of structural and social role requirements guide individual experiences in both the societal and family levels. The structural approach places less emphasis on the socialization of gender-typed behaviors and traits; rather, it examines the common positions occupied by certain groups within the existing social structure. For example, researchers found that occupational roles outweighed gender information as a strong determinant of traits assigned men and women.

Stereotyping is another cognitive process by which gendered beliefs become perpetuated and solidified. Stereotyping presents an avenue through which information becomes cognitively categorized in order for

individuals to process unmanageable amounts of complex information. The formation of stereotypes allows people to form manageable categories based on similarities in which newly encountered people and objects fit. Although stereotypes represent highly efficient modes of organizing information, they often become generalized and oversimplified. Stereotypes represent widely shared and socially validated beliefs or cognitions regarding whole groups of people that become applied to individual members of that group. Stereotypes differ from prejudice, as prejudice refers to an attitude toward a group that often manifests itself as hostile and negative. Therefore, a stereotype represents a cognition or thought, whereas prejudice represents an attitude and often a behavior brought about by that attitude. The cognitive process and the inherent danger of stereotyping applies to gender in addition to ethnicity, socioeconomic class, and religion, for example.

2. GENDER ACQUISITION IN CHILDHOOD

Gender socialization, or the process of learning gender roles, begins at the onset of human life and occurs through cultural, structural, and cognitive avenues. The first statements at the birth of a child include references to the sex of the infant. Parents and friends alike agonize over the purchase of blue or pink gift items according to the sex of the newborn. Beyond the acquisition of items that are "blue for boys" and "pink for girls" and the inherent information presented by these items, infants encounter different treatment based on their sex. Adults, whether knowingly or not, handle infant girls more gently than infant boys, and comment upon the delicate features of girls but on the strong features of male infants. For example, as soon as 24 hours following the birth of a child, mothers and fathers describe female infants as finer featured, softer, smaller, and more inattentive as compared to male infants. This occurs despite the lack of significant differences between female and male infants in both birth length and weight. This difference in treatment occurs even prior to children's acquisition of language and before they become cognitively aware of gender-appropriate behavior.

The acquisition of gender identity becomes apparent in children by the age of three. Social learning theory offers an explanation for the development of gender roles and gender-typed preferences at such an early age. It suggests that children learn gender-appropriate behavior through the observation of others. This occurs through observational learning in which the child watches the behavior of others, parents and other adults alike, and mimics their behavior. In addition, adults reinforce children for acceptable gendered behavior and punish them for violating gendered expectations. For example, if a young boy begins to play with his sister's doll, his parents may look at him disapprovingly or even remove the doll from his play area. In addition to experiencing reward contingencies directly, children witness vicarious reinforcement. Vicarious reinforcement refers to the observed reinforcement or punishment of parents, friends, teachers, and television characters for their gender-consistent or -inconsistent behavior, which children imitate depending on the anticipated reward. The reinforcement of observed others sends messages to the child regarding the acceptability of the observed individual's actions. Therefore, children receive information regarding not only the actual mechanics of behaviors but also the social value in imitating them.

The process of gender socialization also takes place bidirectionally, with children participating in their own socialization. Children cognitively categorize the information they actively learn through the reinforcement and modeling of gender-appropriate behavior. Cognitive-developmental theory suggests that children desire to display prescribed gender-appropriate behaviors. While reward contingencies shape some gendered behavior, children also form their own preferences regarding toys, activities, and same-sex individuals. In a cyclical fashion, social and cognitive processes maintain these gender preferences, and both encourage the reinforcement and performance of gender-consistent behavior. In turn, the performance of appropriate gendered behaviors increases the reinforcement of gendered activities and cognitions.

Children often prefer to wear gender-appropriate clothing and play with same-sex peers. Often this occurs due to children's' own inability to remember or categorize same-sex behaviors and traits correctly. Therefore, children use general in-group (individuals perceived as belonging to the child's own group) and out-group (individuals perceived as not belonging to the child's own group) favoritisms to guide their own behaviors. For example, a girl may express the belief that her same-sex peers are more friendly, honest, and humorous than the same-sex boys group. However, the boys feel the same way toward their same-sex peers and feel negatively toward the opposing group of girls.

Child peer preference also occurs due to the propensity of parents and child-care facilities to segregate children into same-sex play groups based on gender-typed activities. These activities (such as sports or group projects) encourage bonding among their participants and the formation of in-group and out-group stereotyping and grants children a venue for developing more same-sex relationships. In addition, children, like adults, pay attention to and remember behavior consistent with their gender stereotypes. Conversely, gender-inconsistent behavior passes unnoticed and unassimilated into the child's general gender schema. As a result, children learn that boys and girls normally spend time apart and in different activities.

3. THE HARMFUL EFFECTS OF GENDER STEREOTYPING

Due to the classification of gender into "feminine" and "masculine" qualities, traits, behavior, or activities, gender becomes prone to stereotyping. Due to the salience and visibility of sex, it becomes the basis for cognitive categorization (age, race, and ethnicity offer some other examples), and a number of gender stereotypes become founded on the perception of an individual as a member of the male or female sex. Whether stereotypes actually reflect differences between men and women remains debatable. Biology, the environment, or an interaction of the two may actually be responsible for differences in males and females. However, applying stereotypes to individuals presents some danger, despite any apparent truth at the group level. The deeply engrained and subtle nature of gender-associated processes causes most individuals to remain unaware of their impact on behaviors, attitudes, cognitions, language, interpersonal relations, and the structure of social institutions.

3.1. Sex Discrimination and Sexism

Gender stereotyping becomes problematic, especially for women, considering that judgments of female gender are formed in regard to ambiguous and intangible traits such as an individual's "congeniality," "warmth," or "interpersonal communication skills." Research conducted during the 1970s and 1980s indicated that women were portrayed as more expressive and communal. Men were viewed as more competent, agentic, and instrumental. In addition, these gendered differences in beliefs regarding personality traits superseded beliefs regarding appropriate roles and occupations.

Stereotypes of women also differ based on race, class, or ethnicity. For example, traits such as "warm," "vain," "dependent," and "intelligent" were frequently applied to White women, while traits such as "dirty," "hostile," "emotional," and "superstitious" were generally applied to Black women. In addition, people view middle-class women as more "intelligent," "ambitious," "competent," and "happy." Lower class women became viewed as "talkative," "irresponsible," "dirty," and "dependent." Stereotypes of women become especially hazardous considering sex discrimination and sexism.

3.1.1. Sex Discrimination

In general, the term sex discrimination refers to the harmful and unequal treatment of individuals based on their sex. Sex discrimination occurs as a function of cognitive internal processes coupled with the power men exert in invoking anti-female prejudices in everyday practices, resulting in the denial of the same privileges enjoyed by men. Women remain the frequent targets for sex discrimination, which produces greater disadvantage for them. While men experience discrimination based on other social variables (religion, age, race, or sexual preference), their gender neutralizes the experience of these sources of discrimination. However, being female remains the largest predictor of discrimination over any other variable because it is interlaced with other social variables (such as age, race, social class, and ethnicity), which produces a greater disadvantage.

3.1.2. Sexism

The experience, impact, and severity of sexism varies depending on its intentionality, visibility, and ease of documentation. Blatant sexism occurs when women face transparently harmful and unequal treatment. This type of sexism includes physical violence, sexist jokes, degrading language, sexual harassment, and unequal treatment under the law or in employment. Less apparent, subtle sexism remains less visible to others and the targeted individual. More often it lacks the focus of other forms of sexism in its intention, maliciousness, and manipulativeness. This type of sexism might occur when, for example, a male friend jokes about women always leaving to go to the bathroom together. The friend's comment lacks the harmful intention, manipulativeness, or maliciousness of other forms of sexism. This type of sexism can be easily

documented; however, it often passes unreported due to the lack of a perceived threat.

Sexism in the form of deliberate behavior designed to sabotage women's efforts socially or in the work place represents an intentional, hidden, and frequently hostile attempt to thwart women and ensure their failure. For example, a man may go out of his way to slash a woman's tires, making her unable to arrive at work on time, or he may erase the content of her computer prior to an important presentation. This type of sexism presents difficulties in documentation, as the strategies used remain easily denied.

3.2. Suggestions for Minimizing Harmful Stereotypes

The danger of applying gender stereotypes negatively through sex discrimination or through blatant, subtle, or covert sexism raises the question, How do parents and other adults minimize gender stereotyping? Some parents believe that by avoiding the imposition of gender role constraints and pressures their children will develop more fully. There exist different reasons for parental focus on reducing gender constraints on their children, as well as implications of gender role research on parenting practices. Parents give a number of reasons for their desire to decrease their children's exposure to gender-stereotyped roles. First, many parents believe that androgynous characteristics and gender flexibility contribute to mental health. Second, many of these parents believe that the attainment of gender equality depends on the debasement of gender roles in the separate spheres of influence traditionally inhabited by females and males (e.g., the home as the women's sphere of influence and the workplace as the man's sphere of influence). Third, they believe that men and women remain more similar than different. Finally, these parents also believe that traits belonging to both genders increase success in today's technologically savvy world. Parents with such goals sometimes find their efforts frustrated, and psychological evaluations of parenting practices oriented toward decreasing gender-stereotyped cognitions and behaviors remain mixed. This is discussed in the article on gender role development in this section.

Parents who desire to minimize the effects of cultural pressures on gender role identity formation can take several steps. First, to take advantage of children's propensity to display interest in activities preferred by same-sex individuals, parents can give examples to their children of same-sex individuals participating in those activities. For example, if parents want their daughter to participate in track and field competitions, they should show their child pictures of a female world-class athlete winning an Olympic medal. Second, parents should present their children with a range of activities and playthings with the hope that they prefer some non-gendered objects and activities. In addition, parents should request that other family members give gifts intended for both sexes, or those intended for the other sex, and encourage their use equally. Third, due to the powerful effect of parents as role models, parents should attempt to display gender-inconsistent behavior as often as possible. Parents should also seek out other role models, such as teachers, who value nontraditional gender roles, refrain from segregating groups on the basis of gender, and attempt to minimize children's exposure to traditional gender roles through television and literature. Fourth, parents should give their children gender-neutral toys and encourage their participation in activities that both sexes enjoy, thereby decreasing the amount of time children spend in stereotypic gender play. Fifth, parents should remain consistent regarding children's exposure to gender-inconsistent and gender-neutral material and behavior. Often this includes refraining from discouraging gender-inconsistent behavior in their children, such as a boy playing "dress-up." Finally, as children approach middle childhood (ages 9 to 12). they become more cognitively developed. By teaching children critical thinking skills and challenging their dogmatic gender cognitions, parents may successfully counter the effects of gender stereotypes.

4. GENDER'S SEXUAL COMPONENT

Gender roles manifest themselves culturally, structurally, and cognitively in many spheres of human interaction, including sexual interaction. Sexual gender role behavior begins in childhood and develops through life. Gendered sexual behavior becomes learned early in life through biological and social reinforcements ("nature" versus "nurture"). These biological and social reinforcements occur differently for males and females and continue throughout the life cycle.

Early childhood sexual behavior begins as children explore their surroundings and themselves. Children experience pleasure through contact with their genitalia, and this pleasure serves as conditioning for future self-stimulation. Genital pleasure also occurs accidentally through the child's contact with objects that create

stimulating responses. Through exploration children learn about the pleasure that results from masturbation. These biological processes serve a purpose for the developing child, both providing the child with sexual reflexes and assisting in the development of neural pathways in the brain necessary for the experience of sexual pleasure. However, the responses of parents and caretakers to young children's self-initiated stimulation determine the ways in which children feel about sexual behavior and their own sexuality.

Parents differ in their responses to their child's self-stimulation. Some parents adopt a normative view regarding their child's exploration and self-touching. Parents who allow their children to explore their sexuality allow positive Pavlovian conditioning and decrease sexual inhibitions. However, some parents regard self-pleasuring behavior as abnormal, and according to cultural beliefs as dirty or sinful, and often use punishment in response (e.g., spankings, harsh tones, or physical reprimand). Unfortunately, children receiving harsh punishment for self-touching repress their masturbation and experience long-lasting inhibitions. Many parents allow some forms of self-touching but punish others, thereby inhibiting and punishing their child.

Young children encounter many different double standards in regard to sexuality and sexual pleasure. Baldwin and Baldwin, in their article on sexual behavior in this section, discuss parents' unequal treatment of sexual behaviors exhibited by their children. First, parents punish girls for masturbation more often than boys, leaving boys with a greater impression of acceptance of sexual self-stimulation. This causes girls and women to become more uncomfortable with exploring different forms of self-pleasuring. Second, even the language used in naming and referencing of children's genitalia gives evidence to this double standard. While boys genitalia becomes referred to as "pee-pee" or "wee-wee," parents give socially detailed information to their girls in the names they provide, such as "treasure box, to open after marriage." Parents often neglect to inform their children about the uses of their genital organs, especially their use in pleasurable activities.

Throughout their childhood, children learn about sexuality from information gathered through observation of parents, other adults, television, film, and literature. Contrary to adult beliefs, children often remain more attentive to sexual information and learn from observation. Parents often assume that children remain oblivious to the sexual content and innuendos present in media and adult behavior. However, considering children's attention to the abundance of sexually suggestive material available to them, they are more exposed to sexually explicit material than their parents might think. Some concerned parents offer their children explanations of sexual material, thereby opening a dialog and equipping their children to better process sexual information and to incorporate the parent's values into this process. However, some parents fail to offer advice regarding sexually explicit material and risk their children accepting sexual violence as normal and natural. For example, while sexual harassment and rape represent a manifestation of power inequality between men and women in which men view women as objects designed to fill their sexual desires, learned gender stereotypes and sexual scripts contribute to the view of men as the initiators of sexual contact and of women as the gatekeepers who rebuff sexual contact.

Parental lessons taught to children regarding their sexual self-pleasure and their observations carry social implications for adolescent and adult sexual behavior. Parents tend to offer girls more guidance regarding sexually explicit media as compared to boys, possibly due to the fear of possible detriment to girls resulting from sexually transmitted infections, pregnancy, and internal injuries resulting from forcible sexual activity. Parents, adults, and media messages therefore convey a message of responsible sexual behavior for girls, but neglect to focus this message on boys. This message becomes especially potent during adolescence, when secondary sex characteristics emerge, causing males and females to appear more adult-like. Baldwin and Baldwin suggest that biological differences in males (e.g., their more visible and often spontaneous sexual arousal and orgasms, and the more visible changes in outward genital appearance) cause them to become more curious and interested in sexual activity at an earlier age. However, this level of curiosity and interest differ depending on the inhibitions and positive associations formed in childhood through interactions with parents, other adults, and media influences.

The interaction of nature and nurture during adolescence creates gender-related sexual behavior differences. Adolescent females, for example, become more concerned regarding possible pregnancy, abortion, sexually transmitted infections (STIs), and sexually transmitted diseases (STDs). In addition, certain social stigmas become evident during adolescence. For example, adolescent females become labeled as "sluts," "easy," or "loose" if they participate in sexual behavior with a number of partners. However, male adolescents receive the label of "stud," which often improves their social status or popularity, if they participate in sexual

behavior with many females. As a result, female adolescents develop more inhibitions regarding sexual behavior, and they tend to become more concerned about the number of partners they engage in sexual behavior with.

The social context of double standards in sexual behavior also contributes to the actual content of sexual activity in adolescence and often in adulthood. Adolescent females often take passive roles in sexual encounters, due to the belief that "nice girls" do not initiate or aggressively pursue sexual pleasure. In addition, females tend to feel uncomfortable and worry about how their partner perceives them if they seem too eager or display too much sexual knowledge. Conversely, actively pursuing sexual contact and obtaining vast amounts of sexual experience become the expectation for male adolescents. In addition, men often state their motivation for seeking sexual activity as oriented toward sexual pleasure or enjoyment, whereas women view their motivation for sex as the desire for love and commitment. The biological differences between men and women fail to offer any support for the differences between male and female sexual behavior. Rather, sexual behavior differences may result from the role expectations conferred upon both men and women.

4.1. Cultural Views of Sexuality

In a larger cultural context, sexual behaviors and values remain important aspects of all societies. Societies differ in the various meanings attributed to sexual intimacy according to their sexual permissiveness, the expression of their sexual values, views regarding contraception and birth control, and the health issues involved. Some broad generalities exist within this infinite variety and the individual differences existing within each culture. Cultural, socioeconomic, historical, global, and technological contexts determine the commonalities and differences in beliefs and behaviors occurring in societies. The large amount of information available to people through globalization and technology allow them to engage in social comparison and augment their understanding of what is natural and normal, thereby shaping their attitudes toward sexual issues and the acceptability of different sexual behaviors.

The permissiveness or restrictiveness of the culture in which an individual lives determines sexual behavior for both genders. A culture's technological advancement, religious orientation, socioeconomic importance placed on marriage, and significance placed on reproduction influence the degree to which it becomes permissive and accepting of sexual behavior. Positive attitudes toward sexual development as part of healthy social adjustment characterize societies with permissive attitudes toward sexual behavior. Permissive cultures place less restrictions on sexual activities and provide children with sexual information (e.g., through ritual or institutional practice, or through exposure to nudity). Restrictive societies limit exposure to sexual information and attempt to repress sexual expression. These societies often view sexual behavior as dangerous, sinful, or harmful. This view toward sexual behavior often stems from societal values regarding the expected goals of its members. For example, parents in the United States often expect their children to obtain college degrees and marry at older ages. The possible risks of participating in sexually intimate activity (e.g., pregnancy or the contraction of AIDS) run counter to the goals of educational success and economic prosperity. Sexual activity, therefore, becomes risky in these cultures due to the possible threat to an individual's social status, physical health, or psychological well-being. In restrictive cultures, abstinence prior to marriage is stressed, often for women more than for men.

The risks of sexual activity often dictate cultural views of sexual activity prior to marriage. Agrarian and hunter–gatherer cultures often perceive premarital sexual intimacy negatively due to the lack of availability of birth control options. In contrast, industrialized societies increasingly view premarital sexual activity more permissively due to advances in birth control technology, the increasing time gap between sexual maturity and marriage, and rising economic equality between men and women. Societies also differ in their permissiveness of other forms of sexual expression, such as homosexuality and transgenderism.

4.2. Homosexuality

As with heterosexual sexual behavior, cultures display either permissiveness or restrictiveness regarding homosexual sexual behavior according to the sociohistorical and socioeconomic contexts. Generally, homosexuality refers to sexual attraction to members of an individual's own sex. Cultural differences in the definition of "sex" and the meanings attributed to different sexual behaviors between people of the same sex often lead to contrasting ideas regarding what homosexual behavior entails. Cultures with a narrow definition of sex, such as consisting solely of penile–vaginal intercourse, may accept a great deal of same-sex intimate behavior.

Gender role perception remains linked to conceptualization of homosexuality in many societies. In some

cultures, men participate in homosexual relationships yet refrain from identifying themselves as homosexual. The roles that these men undertake during homosexual activity assist them in avoiding the stigma that accompanies this behavior. These men may remain considered heterosexual due to the adherence to the masculine norms of their gender. For example, males in prison may participate in homosexual sexual activity; however, because they maintain "masculine" gender characteristics, may escape the social label of homosexual.

In other cultures, such as the United States, homosexuality is defined less according to gender roles, but rather becomes perceived as an intrinsic aspect of the individual. In cultures ascribing homosexuality as part of the individual, individuals experience flexibility in the sexual roles they assume. Increasingly, many cultures are adopting Western views regarding homosexuality, especially due to the proliferation of scientific and political information regarding a possible genetic basis for homosexuality. Due to the Puritanical nature of Western society and its history of discrimination directed toward homosexual individuals (including hate crimes), the shift in values creates a difficult atmosphere for those with other sexual orientations.

4.3. Transgendered Individuals

Some individuals adopt the gender roles of the opposite sex. These transgendered individuals prefer the gender roles or identities of the opposite sex and either continue to behave in accordance to the roles granted them by society or outwardly display the identity and roles they perceive as their true gender. Some transgendered individuals choose to adopt the dress and mannerisms of the other sex, or pursue sex reassignment surgery. Transgenderism does not necessarily apply to cross-dressing, because the motivation for cross-dressing does not always entail a strong preference for opposite gender roles and identities.

4.4. Homophobia

The tendency of other cultures to adopt Western views regarding homosexuality becomes increasingly problematic due to the reflection of Puritan attitudes regarding sexual behavior and gender roles in the United States. Individuals fundamentally threatened by homosexual attractions and relationships claim that traditional morals and family values become eroded by the acceptance of homosexuality. These attitudes rest at the root of discrimination based on sexual orientation. Individuals identifying themselves as homosexual, or even appearing to display homosexual "traits," experience discrimination (e.g., in the workplace, regarding club membership, in school attendance, in obtaining housing and credit, and in obtaining services at places of business) despite the outlaw of discrimination and violence directed toward individuals based on sexual orientation. Discrimination based on sexual orientation, while illegal, occurs frequently in Western society.

On the individual level, part of the discrimination against homosexuals, bisexuals, lesbians, gays, transsexuals, or transgendered individuals occurs due to homophobia, or anti-gay attitudes or behaviors. The term homophobia implies an intense fear of homosexuals (being touched, looked at, or approached by a homosexual) due to the use of the Greek word "phobia," meaning fear. However, while some individuals exhibit an intense fear of homosexuals, the term best describes a heterosexual bias in which heterosexuality is viewed as the norm and superior compared to homosexuality. Often this bias causes individuals to ignore other forms of attraction and attitudes, focusing solely on heterosexual attractions and attitudes. The same prejudices found to contribute to the discrimination faced by women also apply to the discrimination faced by homosexual individuals. Homosexual individuals become evaluated or judged as a group toward which many members of society express negative attitudes and beliefs. Heterosexuals who adhere to a more "traditional" ideology of family and gender, who hold strong religious beliefs, and who maintain friendships with people expressing anti-homosexual attitudes tend to express stronger anti-homosexual beliefs and attitudes. Heterosexuals who experience more contact with homosexual individuals express fewer negative attitudes toward homosexuality as compared to those with little contact with homosexual individuals.

Homophobia affects homosexual individuals' social well-being as well as their emotional and psychological well-being. Internalized homophobia occurs when homosexual, lesbian, gay, bisexual, transgendered, or transsexual individuals feel negatively about themselves based on their sexual orientation. Homosexual individuals also risk adjustment difficulties arising not only from internalized homophobia, but also from the social stigma attached to homosexuality. In addition, very pronounced fears regarding the AIDS pandemic and its labeling as a "gay" disease caused many homosexual individuals to resist their sexual orientations in order to avoid contraction of the virus. Although the stigma of AIDS as confined to the homosexual and

intravenous needle-using population abated and this belief became replaced by more informed views, many homosexual individuals still experience the stress of anti-gay sentiment.

Homosexual men and women experience high levels of stress due to their marginalization in society. Historically, the mental health field confronted homosexuality as a form of psychopathology; however, this view evolved somewhat and many mental health professionals remain supportive of gay and lesbian identities. Although both the mental health field and society are moving toward a more accepting view of homosexuality, greater attention is needed in addressing the social stressors contributing to the developmental difficulties faced by homosexual individuals.

5. MARRIAGE AND PARENTHOOD

5.1. Marriage

Marriage represents another social context in which gender plays a large part. In most societies, marriage remains an antecedent of the nuclear family, consisting primarily of a husband, wife, and children. The larger, extended family consists of those related by blood to the members of the nuclear family, for example, the grandparents, aunts, uncles, and cousins. The sociohistorical context of the family frames the gender roles occurring within its influence.

During most of human history, people perceived marriage as essential due to the needs of both men and women. The labor of caring for and maintaining a family became shared by both the man and the woman, usually with the man undertaking the role of the hunter and breadwinner and the woman being responsible for the care of children. Both men and women needed the product of the other's labor, not only for their individual survival, but also for the survival of the next generation. In addition, in many cultures marriage remains beneficial in regulating sexuality, providing legal ties between women, men, and their children, and creating economic and social stability.

Women receive the greatest amount of social pressure to marry. The importance of the family remains paramount in most modern societies, despite the existence of many opportunities for women to work and participate in social life, politics, and religion. From a very young age, girls receive education regarding their eventual role as wife and mother. Women who remain unmarried, or elect to forgo marriage, face social sigma (e.g., they receive the name "spinster") and pity unless they take on a socially accepted role such as a nun. As a result of the social stigma attached to remaining unmarried, many women experience a feeling of inadequacy or deprivation. Consequently, families attempt to ensure marriage for their daughters (in some cultures, with arraigned marriages).

For many married women, large segments of daily life remain occupied by the roles of wife and mother, with family obligations at the forefront, despite other obligations. For example, many women in the United States work and care for their families, often maintaining the role of primary caretaker for their children as well as full-time employment. In most Western societies, household maintenance remains unequally allocated between husband and wife, often with the wife assuming the lion's share of daily domestic activities such as cooking, cleaning, laundry, and child care. Evidently, the increase of time-saving technology, the changing social landscape in regarding the increase of women in the workplace, and the rise of gender equality ideals have failed to fully evaporate traditional gender roles.

Evidence of the asymmetry in traditional marriages occurs in various facets of marriage customs, including the assumption of the husband's last name. This also occurs in the conceptualization of marriage in regard to occupations. The expectation of the wife as a "junior partner" in a man's occupation occurs presently in many societies. For example, the wife of a baker becomes the "baker's wife," or the wife of a wealthy banker becomes the "banker's wife." The wife becomes incorporated into the husband's occupation, as an extension of himself. The "baker's wife" may not appear on the payroll; however, she participates in the daily activities of the bakery. The "banker's wife" may appear as the social face of her husband's occupation, contributing to charities and attending social functions. Wives of politicians provide another example; society expects them to place their aspirations aside in order to participate and contribute to their husbands' endeavors. To a large extent, societal expectations of gender roles in marriage dictate the experiences of married women.

In some societies, wives remain subservient to their husbands. In fact, in most parts of the world women's rights and entitlements remain secondary to those of their husbands. A woman may lack control over finances, lack legal rights to inheritance, be considered the property of her husband, and in some societies become the ward of her children in the event of her husband's death. In past extreme cases, women in India were expected to burn themselves alive following

the death of their husband (suttee). The Indian government outlawed the practice, and the actual current incidence remains questionable. The *Code of Manu* (a list of life rules well established by 200 B.C.E.) stated, "She must on the death of her husband allow herself to be burnt alive on the same funeral pyre; then everybody will praise her virtue" (Noss & Noss, 1984; p. 181). Although suttee represents an extreme situation, many women still face societal judgment based on their ability to live up to gender role stereotypes such as maintaining a cleanly and ordered household.

In the past, societies expected individuals to marry for greater financial stability, to reproduce through engaging in socially approved forms of sexual activity, and to form family alliances. In many societies, marriage forms the foundation of families and basis of society. However, other political, social, and economic institutions in industrial urban societies often debase the perceived social function of marriage. Improved technology and economic organization decrease the need for the division of labor based on gender. However, the perception of marriage and the inherent expectations that accompany the institution contribute to the perpetuation of gender role stereotypes and expectations. In addition, due to continued inequalities for women, such as unequal pay, women still maintain strong financial and social motivations for marriage. Same-sex marriages bear mentioning here. While some countries allow same-sex marriages, homosexual couples in the United States struggle for the permission to marry. (Massachusetts is the only state in the United States currently permitting same-sex marriage.) Same-sex couples desire the social, financial, and legal security afforded to heterosexual couples. Therefore, the motivations behind same-sex couples' desire to marry remain the same as the motivation behind heterosexual marriages.

5.2. Parenthood

While many couples choose to marry later and improve their quality of life prior to the birth of their first child, the decision to have children often accompanies a shift toward traditional gender roles in the family. This occurs in part because of the perception of motherhood as the ultimate fulfillment of womanhood. For example, until recently, pregnant women left the workforce in anticipation of children and became full-time homemakers and mothers.

The expectation of women assuming the sole role of motherhood following childbirth grew out of the view that mothers maintained the most important influence over the child's development. However, further research suggested that children with fathers who play an active role in child rearing develop higher intellectual capacity and higher levels of empathy toward their peers. This affects fathers as well; they display higher self-esteem in general, greater satisfaction in their role as a father, and higher confidence in their parenting abilities. Fathers sharing in household responsibilities and childcare also experience greater marital stability.

Despite the benefits of father participation in household and childcare responsibilities, the mother remains more likely to remain home with the child after birth, at least part-time. The mother, therefore, participates more fully in meeting the child's needs, primarily because she remains in the household while the father continues to work. Often this causes the woman to undertake the management of the family's emotional life (including worrying about the state of the marital relationship), the child's everyday needs, and social relations with extended family members.

Yet women in Western societies often return to work, either following the termination of their maternity leave or after their children depart the home. However, parenting demands remain stronger for women. Even with provisions for childcare, working mothers often experience conflicting demands of their jobs and their children, sometimes to the detriment of their mental health. This conflict often becomes exacerbated in the case of single-parent households.

5.3. Single Parenthood

Single parents face a different set of difficulties, as they assume both gender roles. Single mothers and fathers may experience role overload as they attempt to ensure their family's economic survival in addition to managing the household and finances. Employers are often insensitive to the conflicting needs of work and home, expecting full commitment from their employees. Single-parent difficulties become compounded due to the antiquated design of the workplace, which best suits men with wives who manage the household. These issues become especially difficult for single mothers, who face the same discrimination experienced by other working women; however, the situation becomes more desperate considering their status as the provider for the family.

While some single parents encounter social difficulties fitting in with married couples and encounter difficulties in dating, as the number of single parents increases, more opportunities for forming social support networks

are arising. In addition, some divorced women express a sense of accomplishment and fulfillment, while many single mothers enjoy their autonomy and independence. Single fathers also make caring for their children an important priority. Despite the difficulties faced by single parents, single parenthood offers the custodial parent an opportunity to experience both gender roles.

5.4. Gay and Lesbian Parents

Gay and lesbian parents, either single or in committed relationships, find themselves in a unique position. These parents face many problems resulting from the social stereotypes regarding their abilities as role models for children. Despite the concerns put forth by society, many homosexual parents have children from previous heterosexual relationships, adopt children, obtain the assistance of a surrogate, or undergo in vitro fertilization. Contrary to public concerns regarding gay and lesbian parents' abilities to raise well-adjusted heterosexual children, studies indicate that children raised in households with gay and lesbian parents enjoy outcomes similar to children raised in households with heterosexual parents. The raising of children by homosexual couples does not increase the likelihood of children becoming homosexual later in life. While these results show promise, homosexual parents still struggle to protect their children from the negative public opinion of gay and lesbian parenthood.

6. GENDER AND OLD AGE

6.1. Gender and Medical Concerns

Men and women experience the aging process differently, both physically and psychologically. Both medical and psychological literature contain gender bias, as researchers often falsely assume either equalities or differences between the sexes. Many researchers conduct their research using only male samples and then generalize the results to females. For example, researchers receive promising results when using vitamin therapy to decrease hair loss in men and advocate the therapy for use by both men and women. However, the vitamin therapy may not work as well, or at all, in women. While this is a fictitious example, some researchers approve treatments for women regardless of their biological differences, sometimes with detrimental results. In addition, researchers also overlook the incidence of diseases, defects, and illnesses by labeling them as either "men's" or "women's." For example, many prevention strategies target women for osteoporosis. Television commercials advertising calcium consistently refer to women's bone loss and the necessity of increased calcium consumption. However, although women have a higher rate of osteoporosis (one in two women experiences osteoporosis), men also experience bone loss (one in eight experiences osteoporosis). This may lead men to ignore the signs of bone loss or seek treatment later after serious damage occurs.

In addition to the research and prevention strategies for illnesses and disease, men and women experience different illnesses based on their biology. Women have different medical concerns throughout their lifespan (during pregnancy, menstruation, contraception, and menopause, for example). In addition, men and women differ in the incidence and severity of mental illness, due to their differing physiology and environmental stressors. Researchers and other health professionals now seem to recognize that both similarities and differences exist between females and males. In fact, the recognition of women's health as a distinct entity attests to the growing acceptance of quantitative and qualitative differences in the effects of health-related issues on both men and women.

6.2. Gender and the Psychological Experience of Aging

Women and men also experience aging differently psychologically. Women often view aging negatively as the loss of youth characterized by the loss of attractiveness, and the loss of the ability to produce and nurture children. Men, however, view the graying of hair and wrinkling of skin as adding character. The appearance of age does not threaten the traditional gender stereotypes of men as competent, autonomous, and self-controlled; in fact, the appearance of age often serves to "add character." This creates a double standard: as women become viewed as less attractive due to the appearance of age, men become viewed as more attractive for the very same age indicators (e.g., some extra pounds, wrinkles, and gray hair). Youth-oriented society plays a large part in reinforcing views of aging and its experience, especially for women. One study found that the promise of "youth, youthful appearance, or the energy to act youthful" appeared in 57% of monitored television advertisements. However, senior citizens in

these same advertisements dealt with loose dentures, constipation, and backaches.

6.3. External Perceptions of Aging

The emphasis placed on youth and the appearance of youth in culture also influences the ways in which aging individuals are perceived. Researchers found that older women were perceived as more active in the community, more personally acceptable, less self-centered, more nurturing, and more sensitive as compared to older men. However, an older woman perceived as lacking competence is judged more harshly and is perceived as experiencing cognitive decline.

6.4. Ageism

Because women often outlive their male counterparts and make up the majority of the elderly population, they necessitate greater medical and psychological consideration. Older men usually are cared for by women, making them less susceptible to the effects of ageism. Often women do not share this luxury. The lack of any buffers against ageism causes women to become the subject of more age-related jokes. This becomes especially problematic if the loss of her husband and the constraints of living on a restricted budget require an elderly woman to seek employment. Age discrimination occurs in large part due to the view that women's responsibilities as child caregivers decline and end in middle age, therefore placing them at the end of their responsibilities and accomplishments.

7. FEMINISM

Feminism represents one approach to addressing the necessity of different perspectives on gender, as opposed to a purely male-oriented perspective, in research, theory, and practice. The term feminism refers to a belief in and commitment to the social, legal, and economic equality of women. This includes the belief in the equal value of women as well as their equal rights. The broad-based perspective and the incorporation of different ideologies provide feminism with a unique frame of reference. There exist different kinds of feminism (socialist, radical, cultural, and liberal); however, each viewpoint and theory intersects and maintains commonalities with other theories of feminism.

Socialist feminism focuses on social relationships and the preservation and promotion of male dominance in social institutions. The necessity of change in the economic system remains the primary focus of socialist feminism, as economic and sexist oppression reinforce each other in society. Inequalities stemming from race, socioeconomic class, and ethnicity also become the focus of research under this theory. This type of feminism stresses the unity and integration of all aspects of women's lives in order to produce unified feminist theory. Social feminists desire the restructuring of social relationships and social institutions to create economic equality, and therefore social equality, between men and women.

Radical feminism focuses on men's control over and oppression of women, which remains the primary source of women's oppression economically and socially. This theory advocates for the destruction of existing class and gender inequalities and the creation of a new culture and society founded upon shared power between the sexes. Radical feminists believe that women-centered beliefs and systems remain essential to the end of women's oppression. In contrast to other feminist groups, radical feminists reject all forms of feminism based in male-developed theories. In their view, women must create and develop feminism and root it in their own experiences. Radical feminist research often examines the relation between men's societal power and sexuality and violence.

Cultural feminism emphasizes the qualities and characteristics generally applied to women, such as caring, nurturing, and concern for the needs of others, which society usually devalues and ignores. Cultural feminism focuses on increasing the acknowledgement and honor of women's contributions (usually unpaid) to society and culture.

Liberal feminism, or "equal opportunity feminism," focuses on the promotion of legal and social equality between men and women. This theory advocates for a revision of values, laws, and mores associated with the inequality of women in society. In addition, the concept of males and females as more similar than different remains central to this theory, as liberal feminists believe that women and men will exhibit similar behavior if afforded equal opportunities. Liberal feminists call for equal consideration regardless of gender; they do not seek special privileges. Liberal feminism claims responsibility for revisions in language, media representation of women, and changes in religious practices, which formerly contributed to stereotypical views of women in society and inequality based on gender. Due to the stigma attached to the term feminism, many people avoid classifying themselves as feminists.

However, many individuals express beliefs in gender equality consistent with liberal feminism.

8. SUGGESTIONS FOR APPLIED PSYCHOLOGY

The adoption of feminist ideals regarding gender equality in scientific research, theory, and practice would serve to increase sensitivity regarding the needs of both men and women. These ideals serve to exact change in the cultural, structural, and cognitive views of gender, gender roles, and gender-appropriate behavior. The expansion of ideologies pertaining to gender role development, behavior, and attitudes to include often-neglected alternative views (such as those inherent in homosexuality and transgenderism) might serve to lessen the incidence of the discrimination of historically marginalized groups. In addition, a closer understanding of gender stereotypes and their formation along with social and cultural ideological change might contribute to fewer incidences of medical and psychological problems associated with gender discrimination, sexual harassment, and rape.

See Also the Following Articles

Gender Role Development ■ Homophobia ■ Rape Prevention ■ Risk-Taking and Sexual Behavior ■ Sexual Behavior ■ Sexual Behavior and Culture ■ Sexual Harassment ■ Women's Health

Further Reading

Benokraitis, N. V., & Feagin, J. R. (1995). *Modern sexism: Blatant, subtle, and covert discrimination* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Broverman, I. K., Vogel, S. R., Broverman, D. M., Clarkson, F. E., & Rosenkrantz, P. S. (1972). Sex-role stereotypes: A current appraisal. *Journal of Social Issues*, *28*, 59–78.

Carter, B. (1992). Stonewalling feminism. *Family Therapy Networker*, *16*, 64–69.

Deaux, K., & Emswiller, T. (1974). Explanations of successful performance on sex-linked tasks: What is skill for the male is luck for the female. *Journal of Personality and Social Psychology*, *29*, 80–85.

Deaux, K., & Kite, M. (1993). Gender sterotypes. In F. L. Denmark, & M. A. Paludi (Eds.), *Psychology of women: A handbook of issues and theories* (pp. 107–139). Westport, CT: Greenwood Press.

Demo, D. H. (1992). Parent-child relations: Assessing recent changes. *Journal of Marriage and the Family*, *54*, 104–117.

Denmark, F. L., Rabinowitz, V., & Sechzer, J. (2005). *Engendering psychology: Women and gender revisited*. Boston: Allyn and Bacon.

Gilbert, L. A. (1994). Reclaiming and returning gender to context: Examples from studies of heterosexual dual-earner families. *Psychology of Women Quarterly*, *18*, 539–584.

Harre, R., & Lamb, R. (1986). *The dictionary of personality and social psychology*. Cambridge, MA: MIT Press.

Hochschild, A. (1989). *The second shift: Working parents and the revolution at home*. New York: Viking.

Keshet, J. K. (1980). From separation to stepfamily: A subsystem analysis. *Journal of Family Issues*, *4*, 146–153.

Kimmel, D. C. (1988). Ageism, psychology, and public policy. *American Psychologist*, *43*, 175–178.

Kite, M. E., Deaux, K., & Miele, M. (1991). Stereotypes of young and old: Does age outweigh gender? *Psychology and Aging*, *6*, 19–27.

Lamb, M. (Ed.). (1987). *The father's role: Cross-cultural perspectives*. Hillsdale, NJ: Erlbaum.

Lamb, M., Pleck, J., & Levine, R. (1987). Effects of increased parental involvement on fathers and mothers. In C. Lewis, & M. O'Brien (Eds.), *Reassessing fatherhood* (pp. 108–125). London: Sage.

Lott, B. (1995). Distancing from women: Interpersonal sexist discrimination. In B. Lott, & D. Maluso (Eds.), *The social psychology of interpersonal discrimination* (pp. 12–49). New York: Guilford Press.

Noss, J. B., & Noss, D. S. (1984). *Man's religions* (7th ed.). New York: Macmillan.

O'Connell, A. N., & Rotter, G. (1979). The influence of stimulus age and sex on person perception. *Journal of Gerontology*, *34*, 220–228.

Offen, K. (1988). Defining feminism: A comparative historical approach. *Signs: Journal of Women in Culture and Society*, *14*(11), 119–157.

Palmore, E. (1971). Attitudes toward aging as shown by humor. *Gerontologist*, *11*, 181–186.

Richmond-Abbott, M. (1984). Sex-roles of mothers and children in divorced, single parent families. *Journal of Divorce*, *8*, 34–46.

Rossi, A. (1979). Transition to parenthood. In P. Rose (Ed.), *Socialization and the life cycle* (pp. 132–145). New York: St. Martin's Press.

Rubin, J., Provenzano, F. J., & Luria, Z. (1974). The eye of the beholder: Parents' views on sex of newborns. *American Journal of Orthopsychiatry*, *44*, 512–519.

Rubin, L. (1983). *Intimate stranger: Men and women together*. New York: Harper & Row.

Silverstein, L. B. (1996). Fathering is a feminist issue. *Psychology of Women Quarterly*, *20*, 3–37.

Simon, R. W. (1992). Parental role strain, salience of parental identity and gender differences in psychological distress. *Journal of Health and Social Behavior*, *33*, 25–35.

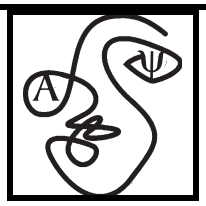
Spence, J. T., Helmreigh, R. L., & Stapp, J. (1974). The personal attributes questionnaire: A measure of sex-role

stereotypes and masculinity/femininity. *JSAS Catalog of Selected Documents in Psychology*, *4*, 43.

Stockard, J., & Johnson, M. M. (1992). *Sex and gender in society*. Englewood Cliffs, NJ: Prentice Hall.

Tong, R. (1989). *Feminist thought: A comprehensive introduction*. Boulder, CO: Westview Press.

West, C., & Zimmerman, D. H. (1987). Going gender. *Gender and Society*, *1*, 125–151.



Gender Role Development

Joan C. Chrisler
Connecticut College, New London, Connecticut, USA

1. Toward a Modern Individualist Society
2. Learning to Conform to One's Gender
3. Acquisition of Gender Identity
4. Active Participation in Gender Socialization
5. Shaping Future Women and Men
Further Reading

GLOSSARY

gender A term used to classify individuals based on psychological and behavioral characteristics, social activities, and cultural expectations that are often based on beliefs about people's sex; the terms girl/woman/feminine and boy/man/masculine are gendered classifications.

gender identity A term used to describe an individual's own belief about his or her gender classification (e.g., "I am a girl," "I am a boy"); by 3 years of age, most children's gender identity is fixed, and in most cases, it matches their biological sex.

gender role Social norms for the behavior of girls/women and boys/men; these are considered to be meta-roles in that they also influence expectations of people's behavior in subroles.

sex A term used to classify individuals based on biological characteristics related to reproductive capacity (e.g., chromosomes, hormone levels, gonads, genitalia); most cultures maintain a binary classification of individuals as either female or male, although some have long recognized the presence of intersex individuals by providing three or more categories of sex.

This article describes psychological, social, and cultural factors that lead to the development of gender roles. It provides advice for parents who wish to encourage their children to be more individualist and less gender stereotyped in their behavior.

1. TOWARD A MODERN INDIVIDUALIST SOCIETY

Parents who describe themselves as modern, liberal, feminist, and/or nontraditional often say that they want to raise children who are free from gender role constraints and, thus, are able to develop into full human individuals who can be "themselves." These parents set that goal for various reasons, including some (or all) of the following: (a) the belief that the flexibility of androgyny is good for people's mental health; (b) the conviction that women and men will never reach political and economic equality so long as traditional gender roles maintain the expectation of separate spheres (i.e., that women belong in the private sphere and should exercise power in the home, whereas men belong in the public sphere and should exercise power in business and government); (c) the belief that women and men are more alike than they are different and that, therefore, similarities (rather than differences) should be emphasized; and (d) the fact that technology (e.g., contraceptives, power tools, computers) has changed modern life to such an extent that the specialized talents (e.g., nurturance, physical strength) that gender

©2004 Elsevier Inc.
All rights reserved.

roles describe are no longer essential for success in modern life and that, instead, both women and men need combinations of feminine and masculine traits to succeed. These parents tell their children that they can be or do anything they choose, provide them with both traditional and nontraditional toys and games, and try to model androgynous behavior themselves, for example, by deliberately showing their children that both women and men can be both competent and kind.

How successful are these parents in reaching their goal? The results are mixed, and parents find that it is especially difficult to encourage androgyny in boys and in both girls and boys when they are young. The parents become frustrated and confused, and sometimes they abandon their goal. They ask psychologists questions such as the following: "Why are my sons macho and my daughters girly despite my efforts to provide them with all sorts of opportunities? What am I doing wrong?" The answer is that the parents might not be doing anything wrong; in fact, they might be doing everything right. Powerful messages from peers and popular culture can override parents' best efforts.

2. LEARNING TO CONFORM TO ONE'S GENDER

Gender socialization (i.e., the process of learning one's gender role) begins as soon as babies are born. The hospital staff members announce, "It's a boy!" or "It's a girl!" As soon as the newborn is cleaned up and weighed, it is wrapped in either a pink or a blue blanket. Relatives and friends of the new parents rush to the store to buy congratulatory cards, and they find themselves confronted with an array of blue or pink cards containing gendered messages about how active a baby son is and how sweet a baby daughter is. One must search long and carefully to find a unique nongendered card, which might be yellow and green and discuss how much happiness a new baby can bring. Before babies can even begin to understand spoken language, adults and older children greet them with gendered exclamations such as the following. "What a big boy!" "How's my little girl?" "He has such a strong grip!" "She has beautiful eyes!" Adults handle baby girls more gently and play with baby boys more vigorously. Some of these actions are done deliberately because people believe that it is proper to treat boys and girls differently. Other actions are done without thinking about their repercussions. The phrases "big boy" and "sweet girl" just seem to go together because people have heard those phrases uttered so often.

3. ACQUISITION OF GENDER IDENTITY

Children acquire their gender identity by 3 years of age, and they begin to exhibit gender-typed preferences even earlier. In 2002, Poulin-Dubois and colleagues demonstrated that 2-year-olds spent more time looking at photographs of people behaving in gender role-inconsistent ways (e.g., a man putting on makeup) than they spent looking at photographs of people behaving in gender role-consistent ways (e.g., a woman ironing). This suggests that even at that young age, the children found the gender role-inconsistent photographs to be surprising or puzzling. Social learning theory has often been used to explain the development of gender roles and gender-typed preferences. The theory posits that children learn about gender by watching how others behave (i.e., by observational learning or modeling), that children are rewarded by others when they behave according to gendered expectations, and that children are punished by others when they do not (i.e., reinforcement contingencies are applied).

Although parents are the most powerful models that young children observe, they are certainly not children's only important models. Grandparents and other relatives, teachers, coaches, neighbors, religious leaders, television characters, and other children also are influential. In addition to modeling gendered behavior, all of these (with the exception of the fictional characters) are also able to reward children for exhibiting gender role-consistent behavior and to punish them for exhibiting gender role-inconsistent behavior. Rewards take forms such as smiles, nods, hugs, kisses, praise, and gifts. Punishments take forms such as frowns, head shakes, withdrawal of affection, verbal corrections, and even spankings. For example, "Susie" may notice that her older sister has a doll collection, whereas her older brother does not. Susie's grandmother may give her a doll for her birthday and exclaim, "I'm so glad that you like it, sweetheart!" In preschool, Susie may notice the teacher frown when she joins the boys in playing with the transportation toys, and her girlfriends may offer a behavioral correction by calling out, "Susie! Come play dolls with us."

3.1. Media Influences

The media are very strong influences on children's behavior, and they teach lessons of all kinds about the world and people's roles in it. Researchers have found that 3- to 5-year-old North American children watch an average of 15 to 20 hours of television per week. By 7 years of age, this drops to an average of 15 hours or less, mainly because the children are spending more hours per week in school. The youngest children may watch primarily educational programming, such as *Sesame Street* and reruns of *Mister Rogers' Neighborhood,* which tend to emphasize the uniqueness of individual children. However, older children watch primarily cartoons, action/adventure shows, sports, and situation comedies, all of which portray gender role stereotypes. On average, boys watch more television than do girls. Children who watch television learn many things about women/girls and men/boys. They learn that boys are noisy and active, that girls are quiet and passive, that women are victims who need to be rescued by men (heroes) from other men (villains), that boys and girls play with different toys and games, that women and men pursue different hobbies and careers, that girls should be pretty, and that boys should be strong. Children's books and movies contain similar messages, and these messages are repeated frequently as children reread and rewatch their favorite books and videos.

3.2. Division Along Gender Lines

Children's worlds are divided along gender lines. Toy stores are clearly demarcated by gender, with "girls' toys" on one side and "boys' toys" on the other. An adult who enters a toy store and asks a clerk for help in choosing a gift for a 5-year-old will immediately be asked, "Are you shopping for a boy or a girl?" Children who enter the store generally know exactly where to go because they have seen girls and boys modeling toy use on television commercials, and they head for the area of the store that features the "appropriate" toys. If there is any doubt about which aisles are for them, the colors of the toys, the packaging, and sometimes even the walls and shelves of the store will serve as clues, with bold colors (especially blue and red) being for boys and pastel colors (especially pink and purple) being for girls. Children's clothing is similarly color coded and is often decorated with gender-typed emblems, for example, a football or firetruck for boys and a flower or rainbow for girls. Girls' bibs, socks, and hats are often lacey, and stores sell barrettes and headbands for use even with newborn girls who have very little hair. When children start to attend day care, preschool, or other organized activities, they will find themselves addressed as "boys and girls" far more often than as "children," a pattern that will continue through the elementary school years. Although the phrase may be more habitual than deliberate on the part of the adults who use it, it communicates powerfully to children that people come in two "types."

4. ACTIVE PARTICIPATION IN GENDER SOCIALIZATION

Gender role socialization is not a passive process. Once children realize that people come in two types and figure out which type they are, they tend to categorize by gender everything they possibly can. Cognitive developmental theory posits that children want to behave in a gender role-consistent manner; therefore, they begin at an early age to gather information about what boys and girls are supposed to be and do. The modeling and reinforcement contingencies described previously are sources of information for children to categorize. However, external rewards and punishments might not be the only forces that shape gendered behavior. Children's own internal contingences are also at work as they come to state their preferences for gendered toys and activities and even for same-sex people. Thus, children's inner desire to act like girls or boys may be responsible for the dynamics that causes their parents to complain to psychologists as in the following ways: "I prefer to dress my daughter in pants for her warmth and comfort, but she insists on wearing dresses to school." "I want my son to have a variety of toys and games, but all he wants for his birthday are trucks and guns."

Social and cognitive processes help to maintain children's gendered preferences. Children spend a lot of time in gender-segregated play groups, that is, all-boys groups or all-girls groups. This segregation is, in part, both a cause and an effect of children's preference for same-sex peers. This preference is maintained in part by children's tendency to misremember, or miscategorize, people's traits and behavior so that same-sex traits and behavior are more highly valued. For example, boys believe that boys are nicer, smarter, and funnier than girls, and girls believe that girls are nicer, smarter, and funnier than

boys. In-group/Out-group processes also help to maintain gender stereotypes and preferences for same-sex others. Any groups whose members bond together (e.g., gender groups, ethnic groups, sports teams, artificially created groups in a social psychology experiment) look for ways in which to distinguish themselves (the in-group) from others (the out-group). They exaggerate the differences between the in-group and the out-group, form a strong preference for in-group members over out-group members, and believe that in-group members have more variability in traits and behaviors than do out-group members, who are all thought to be alike. Furthermore, once traits and behaviors are categorized by gender, children tend to notice and remember instances that are consistent with gender categories, whereas they might not notice or might forget instances that are inconsistent. These processes may help to explain the logic of a boy who values soccer, thinks that it is a boys' game, and believes that girls do not like soccer even though he has seen some girls play it and knows that not all boys like it.

Gender segregation during childhood also occurs because children are assigned gender-stereotyped tasks (e.g., "Boys, go help your father with the yardwork," "Girls, I want you to help mother with the holiday baking") and because children engage in gendered activities. For example, children learn from older children, as well as from television and books, that girls take ballet lessons and boys play Little League baseball. Of course, one can find boys dancing ballet and girls playing baseball, but they are few in number. Children who choose gender-typed activities will generally find themselves in segregated groups. These activities provide the opportunity to make more same-sex friends and strengthen children's beliefs that it is normal for girls and boys to spend a lot of their time apart. Maccoby's research in 1998 showed that spending so much time with same-sex peers leads boys and girls to develop different play styles. Boys play in larger groups, and their games are often described as rough and competitive. Girls play in smaller, more cooperative groups, and their games are often based on fantasies about becoming attractive (even glamorous) grown-ups. Once these play styles become well established, it is even more unusual for a child to join a cross-sex play group. Boys are more likely than girls to enforce gender-segregated groups; they often actively exclude girls from their games and tease other boys who voluntarily play with girls.

So, what's a parent to do? Is it impossible to raise androgynous children? Should parents give up and accept the status quo? Parents should not give up, but they probably need to be more realistic about what can be achieved, especially at early ages. Until greater social, cultural, and political changes are achieved, parents will face an uphill battle because children will receive dozens of stereotypical messages for every stereotype-inconsistent message parents may provide. However, once children are old enough to have achieved a high level of cognitive development, parent may have greater success.

5. SHAPING FUTURE WOMEN AND MEN

This section contains some advice for parents who wish to raise androgynous children as well as some reasons for these parents to remain hopeful. First, researchers have shown that children show more interest in unfamiliar objects and activities when they are told that the objects and activities are preferred by same-sex people. For example, if parents want their son to take violin lessons, they should show him a picture of Pincus Zuckerman playing his violin and say something about the many famous men who play the violin so expertly. If parents want their daughter to join a soccer team, they should show her a picture of Mia Hamm playing soccer and tell her about the number of girls who are playing in the local youth soccer league. If the first message children receive about something novel is that it is gender appropriate, they may maintain interest in it long enough to find out whether they like it. If they like it, it may become part of the in-group variability discussed previously. A later comment from a peer that "violins are for girls" may be countered by "No, they're not! Lots of boys like to play the violin."

Second, recent studies indicate that both children and adults are more likely to apply gender stereotypes to other people than to themselves. Although this phenomenon is not going to aid in promoting social, cultural, and political change, it does allow individuals to exercise more freedom of choice in their own activities and preferences. Thus, parents should expose their children to a range of toys, games, and activities in hopes that some nongendered objects and activities may become preferred. If children ask only for gendered toys for their birthdays, family members can agree to fulfill some of the requests and to provide some nongendered gifts as well. It would not hurt if the father or grandfather reacts enthusiastically when a boy unwraps a nongendered gift by exclaiming,

"Wow! Candyland [or finger paints]—I loved that when I was a little boy." Researchers have found that boys are less likely to play with a feminine or gender-neutral toy if they think that their fathers would disapprove.

Third, parents should continue to model gender-inconsistent behavior as often as they can. Parents are powerful role models for their children, and children do pay attention to what their parents do, although sociocultural pressures may cause them, at least at younger ages, to deny that gender inconsistency occurs or that it is good. Some studies have shown that children from single-parent homes, where one parent must be both nurturant and agentic, display less gender-stereotypic behavior and fewer gendered preferences than do children from two-parent homes. Also, children who grow up in two-parent homes where the parents are nontraditional in their own gender roles display less gender-stereotypic behavior than do children whose parents are traditional in their role enactment. Furthermore, researchers have found that children in nontraditional preschools spend half as much time in gender-segregated groups as do children in traditional preschools. Parents can also enroll their children in day care centers and preschools where the teachers value and reward nontraditional gender roles, and parents can try to limit the amount of exposure young children have to books and television shows that reinforce traditional gender roles.

Fourth, parents can encourage children to play with gender-neutral toys and to engage in activities that both girls and boys enjoy. Gender-stereotypic thinking is, at least to some extent, context dependent. The more time people spend in gendered activities, the more likely they may be to maintain rigid gender categories. A focus on what all children have in common may help to reduce in-group/out-group divisions. Research on in-group/out-group phenomena suggests that breaking up the existing groups or challenging them to work together to attain a goal will also reduce divisions. Therefore, instead of dividing children in a classroom or at a party into a boys' team and a girls' team for a spelling bee or a game of tag, gender-integrated teams could be formed and given color or animal names (e.g., the reds vs the greens, the dogs vs the horses).

Fifth, parents should be careful not to send contradictory messages about gender expectations, for example, by rewarding traditional behavior one minute and punishing it the next. Parents must remember that they themselves have been exposed to millions of gender-stereotypic messages during their lives; thus, they need to take care to pass on as few as possible. For example, people often interpret the same behavior differently depending on whether it is demonstrated by boys or girls. A child who cries with frustration may be thought to be angry if a boy or upset if a girl, and a parent who applies one or the other of these terms to a child's behavior is teaching a powerful, if unintended, lesson about emotionality. Many parents find it more difficult to be accepting of feminine behavior in their sons than to be accepting of masculine behavior in their daughters because "tomboyism" is not considered by most people to be a problem, whereas "sissy" behavior usually is considered problematic. Parents' concern that boys who play "dress up" will become gay men, or that boys who do not like sports will be unpopular with other boys, can override the encouragement that they give their children to try gender-inconsistent activities.

Finally, as children reach middle childhood (ages 9–12 years), parents can begin to contradict children's dogmatic gendered statements and can actively teach critical thinking skills to counter gender role stereotypes such as in the following ways. "What makes you think that boys shouldn't take violin lessons?" "Oh, look at the skinny girls in this magazine ad. I don't think they look healthy, do you?" It is appropriate for parents to share their values with children who are old enough to understand. This is the time for parents to say, "We want you to dream your own dreams, to follow your heart. We believe that you can succeed at anything you want to try. Whatever you decide to do, we will support you with pride." Children at this stage often become more flexible in their beliefs about, and attitudes toward, gender roles. With encouragement, they may blossom into the androgynous individuals that their parents always hoped they would be.

See Also the Following Articles

Gender and Culture ■ Homosexuality ■ Risk-Taking and Sexual Behavior ■ Sexual Behavior ■ Sexual Behavior and Culture

Further Reading

Furnham, A., & Mak, T. (1999). Sex role stereotyping in television commercials: A review and comparison of 14 studies done on 5 continents over 25 years. *Sex Roles*, *41*, 413–437.

Houston, A. C., & Wright, J. C. (1996). Television and the socialization of young children. In T. M. MacBeth (Ed.),

Tuning in to young viewers: Social science perspectives on television (pp. 37–60). Thousand Oaks, CA: Sage.

Maccoby, E. E. (1998). *The two sexes: Growing up apart, coming together*. Cambridge, MA: Harvard University Press.

Murnen, S. K., Smolak, L., Mills, J. A., & Good, L. (2003). Thin, sexy women and strong, muscular men: Grade-school children's responses to objectified images of women and men. *Sex Roles*, *49*, 427–437.

Poulin-Dubois, D., Serbin, L. A., Eichstedt, J. A., Sen, M. G., & Beissel, C. F. (2002). Men don't put on make-up: Toddlers' knowledge of the gender stereotyping of household activities. *Social Development*, *11*, 166–181.

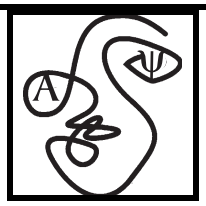
Powlishta, K. K. (1995a). Gender bias in children's perceptions of personality traits. *Sex Roles*, *32*, 17–28.

Powlishta, K. K. (1995b). Intergroup processes in childhood: Social categorization and sex role development. *Developmental Psychology*, *31*, 781–788.

Powlishta, K. K., Sen, M. G., Serbin, L. A., Poulin-Dubois, D., & Eichstedt, J. A. (2001). From infancy to middle childhood: The role of cognitive and social factors in becoming gendered. In R. K. Unger (Ed.), *Handbook of the psychology of women and gender* (pp. 116–132). New York: John Wiley.

Raag, T., & Rackliff, C. L. (1998). Preschoolers' awareness of social expectations of gender: Relationships to toy choices. *Sex Roles*, *38*, 685–700.

Turner-Bowker, D. (1996). Gender stereotyped descriptors in children's books: Does "curious Jane" exist in the literature? *Sex Roles*, *35*, 461–488.



Gerontology, Overview

Peter A. Lichtenberg
Wayne State University, Detroit, Michigan, USA

1. The Merging of Applied and Clinical Geropsychology
2. The Increasing Cultural Heterogeneity of the Aging Population
3. Knowledge Relating to Mental and Behavioral Health Disorders in Older Adults
4. Increasing Use of Applied Geropsychology in Major Life Transitions
5. Principles to Guide Future Activities in Applied Geropsychology
Further Reading

Over the past 15 to 20 years, applied geropsychology has undergone vast and fundamental changes. This article provides an overview of the excellent articles on the topic of applied geropsychology that are included in this encyclopedia, the four major directions in which the field has grown, and the three primary principles that will guide future growth and development. The four major changes that are touched upon by multiple articles in this section are (1) the emergence of applied and clinical geropsychology tools to prevent and treat health problems, (2) the increasing cultural heterogeneity of the aging population, (3) the explosion of knowledge relative to mental and behavioral health disorders in older adults, and (4) the increasing use of applied geropsychology in understanding and assisting with major life transitions.

1. THE MERGING OF APPLIED AND CLINICAL GEROPSYCHOLOGY

In the recent past, applied and clinical geropsychology did not intersect. Applied geropsychologists examined issues relevant to healthy aging and the healthy aged, and the basic process of aging. Clinical geropsychologists examined issues relevant to those older adults needing intervention for aspects of pathological aging, e.g., dementia or depression. The National Institutes of Health (NIH), and in particular the National Institute on Aging (NIA), envisioned the coming together of these seemingly disparate fields of inquiry. By focusing its funding on applications of cognitive aging and other basic knowledge for healthy living, the NIA stimulated the basic understanding of aging with applications of primary prevention and health promotion. At the same time, clinical geropsychology and related disciplines extended their knowledge into the earliest manifestations of pathologies related to aging. For example, the focus on mild cognitive impairment and its relation to dementia has extended early detection from the clinical arena to recognizing that cognitive changes in otherwise healthy aging may be early manifestations of Alzheimer's disease.

The confluence of applied and clinical geropsychology is illustrated in four articles that are included in this section of the encyclopedia. Gould examines aging, cognition, and medication adherence, and how basic cognitive processes of aging are used in designing intervention tools. Ball, Wadley, Vance, and Edwards examine how cognitive skills can be maintained and enhanced. Their

©2004 Elsevier Inc.
All rights reserved.

review of the links between cognitive skills and functional skills and the promising results of intervention trials forecasts a continued growth in the area of cognitive training. Lafleche and Verfaellie analyze the syndrome of amnesia and the problems of clinical patients suffering from one of many disorders that cause amnesia. Cowart's article provides an excellent description of dementia as a clinical syndrome and the use of conceptual models to guide prevention, early detection, and treatment methods. All four articles address the usage of cognitive theories to guide interventions, and demonstrate the ever-increasing convergence between healthy and pathological aging, and the ability to prevent and detect pathology and intervene in clinical syndromes.

2. THE INCREASING CULTURAL HETEROGENEITY OF THE AGING POPULATION

Heterogeneity in aging has traditionally been represented in approaches to healthy versus pathological aging. More recently, it has another, broader meaning in the aging field in terms of the tremendous differences in cultural and ethnic backgrounds among the nation's older adults. Changes in older adult immigration patterns have contributed to a much broader diversity of ethnicity and culture within the elderly in the United States. These dramatic changes are directly examined in the articles by Antonucci and Akiyama and by Brown and Jackson. Antonucci and Akiyama note that cultural experience refers to the interaction between immigration and acceptance. They also emphasize that a life span perspective helps in understanding how individuals become more different from each other as they develop over time, and that the differences among individuals from unique cultural backgrounds may be heightened as they grow older. Brown and Jackson use a life course framework to examine socioeconomic and health disparities between minority groups and European Americans who reside in the United States, and vividly illustrate the heterogeneity in aging between and within minority populations.

3. KNOWLEDGE RELATING TO MENTAL AND BEHAVIORAL HEALTH DISORDERS IN OLDER ADULTS

Applied geropsychology has witnessed an incredible explosion of knowledge in mental and behavioral health relative to epidemiology, measurement tools, treatment and relapse prevention, and the importance of family education. The articles in this section relating to these topics demonstrate the increased sophistication regarding chronic health problems and their assessment and treatment. Related articles on psychotherapy and personality illustrate the growth of knowledge and its application with older adults.

Knight and David's article on psychotherapy presents a conceptual framework that incorporates the similarities of psychotherapy with older and younger adults, while illustrating the unique features of psychotherapy with older adults. Alea, Diehl, and Bluck examine continuity versus the changing nature of personality across older adulthood and the role of emotion in later life. Both frameworks guide their interpretation and thinking about major issues, and provide information that can be used in working with older adults in community and clinical settings.

Depression and anxiety are the two most common forms of mental health disturbance in older age. Late-life depression, which has been the focus of much research in older adults, is examined in the article by Edelstein and Shreve-Neiger. These authors review conceptual models that identify unique aspects of depression in older age and mental and physical health comorbidities. They also examine instruments for use in depression assessment and the interventions used by geropsychologists. Anxiety disorders in older age is examined by Carmin, Mohlman, and Buckley.

Although more prevalent than depression, anxiety lags behind as an area for investigation, possibly due to the confusing array of types of anxiety disorders, each with its own unique diagnostic criteria. Anxiety disorders include a group of syndromes under a single umbrella, such as social phobia, which is clearly distinguished from generalized anxiety disorder. Despite being grouped together, anxiety disorders represent a wide variety of symptom constellations. In contrast, although depression is known to have many different possible etiologies, this disorder is nevertheless defined by a single constellation of symptoms. Consequently, knowledge about anxiety in older adults has lagged behind depression. Carmin, Mohlman, and Buckley examine the contributions that geropsychology is making to the assessment and treatment of anxiety.

Tools from geropsychology have strongly impacted populations such as those with alcohol problems and advanced dementia. People in both groups can easily find themselves isolated and neglected. The unique aspects of alcohol misuse in older adults have been

more clearly identified in recent years. New and broader treatment approaches have also been introduced and used successfully in a variety of settings. Cohen-Mansfield's article on assessment and interventions for behavioral disorders in dementia documents the effective multi-modal approaches developed in geropsychology.

Rybarczyk's article on behavioral medicine issues in late life is a likely precursor to a huge growth of interest in this area of applied geropsychology. His review of pain and sleep disorders and their treatment in older adults reflects the increasing expertise in geriatric rehabilitation that accompanies chronic diseases. Rybarczyk also examines the concept of frailty, which is defined differently by individual researchers. Most work on frailty examines the experience of illness and the adequacy of functioning in the oldest-old. Given the growing numbers of oldest-old, this area will, no doubt, be greatly enhanced by geropsychology in the coming years.

4. INCREASING USE OF APPLIED GEROPSYCHOLOGY IN MAJOR LIFE TRANSITIONS

Applied geropsychology is becoming ever more important in contributing to greater understanding and more effective assistance with major life transitions. This is an exciting area in which society will increasingly look to geropsychology for advances and contributions. Each of the four distinct areas of applied geropsychology that are covered in this encyclopedia reflects increased activity and growing awareness. Individual articles focus on the interrelated topics of caregiving and competency in older persons and end-of-life abuse. Caregiving involves the stressful experience of caring for an older adult with chronic dementia and/or disease, whereas competency focuses on the advanced stages of illness in which an older individual may need help with basic decisions, usually from a caregiver.

Applied geropsychology has developed a much broader understanding of caregiving in the past two decades, including the normative experience of stress, as well as the impact of stress felt by caregivers. Interventions for caregivers continue to proliferate and are badly needed. Burton, Kwak, and Haley examine the ever-growing phenomenon of caregiving for an older relative, which challenges applied geropsychology to develop multiple methods for intervention. Moye discusses the growing knowledge about capacity and competence and reports on research findings and on ethical issues connected to competency. Given the growing knowledge in assessment and measurement, applied geropsychology offers considerable expertise for dealing with issues relating to competency and capacity disputes. The legal system is not well versed in age-related diseases and how they impact specific and broader capacities.

Elder abuse and end-of-life issues in applied geropsychology are two areas of late life in which exploration, to date, has lagged. Elder abuse is an important aspect of caregiving; it is increasingly prevalent and needs greater vigilance. Rodriguez reviews research on elder abuse and examines relevant ethical issues, noting that there are no typical descriptions of abusers or victims of abuse. He also provides clear guidelines for assessment, evaluation, and successful intervention. Allen, Phillips, and Payne review the emerging role of geropsychology in end-of-life issues regarding advanced care planning, palliative and hospice care, and other evolving end-of-life treatment approaches.

5. PRINCIPLES TO GUIDE FUTURE ACTIVITIES IN APPLIED GEROPSYCHOLOGY

The three principles listed below are relevant to all of the articles in the emerging field of applied geropsychology. These principles are also related to and can be used to anticipate future developments.

5.1. Principle 1: Cohort Considerations

Researchers and clinicians must continue to consider the ever-expanding interactions between the physical effects of aging and the psychological effects of cohorts. There are several major developments that are prominent for the first time in research on cohorts. The current cohort of older adults witnessed the great promise of medicine, for example, the discovery of vaccines to prevent polio and other devastating and contagious diseases. These discoveries have contributed to optimism in regard to conquering disease, and to healthy living throughout the lifespan.

Compared to previous cohorts, older adults today are more knowledgeable about the effects of aging and are living longer and healthier lives. Consequently, both clinicians and researchers must be prepared to adapt

when working with more knowledgeable and diverse populations. The growth of the older population, both today and in the future, will impact our society on many levels. Older adults will become a more visible group and will be more often targeted by government, media, and businesses.

In addition to being more numerous, current and near-future cohorts are also characterized by greater diversity in their cultural practices and ethnic backgrounds. The effects of aging are also more diverse than ever before due to these differences. Because there are numerous differences between the current elder cohort and previous and later cohorts, the importance of considering cohort effects, along with the effects of aging, cannot be overstated.

5.2. Principle 2: Conceptual Frameworks

In the previous decade, applied and clinical research developed in parallel and in a relatively noninteractive fashion. There were few applications and limited theoretical understanding (e.g., it was only after 40^{+} years of research that Schaie's longitudinal data were applied to clinical conditions). These new theoretical frameworks are reflected in the contributions of applied geropsychology to this encyclopedia, and are being applied to healthy and pathological aging. Behavioral theory is now used in assisting with treatment adherence among well elders and for decreasing agitation in demented elders. Hopefully, late-life depression theories, such as late-onset and vascular depression, will lead to early detection methods and better depression treatment outcomes.

In the past two decades, clinicians have brought meaningful conceptual frameworks to the study of aging and older adults. Applied geropsychology advances most when theoretical and conceptual frameworks are used in the study of age-related syndromes, and in the application of assessment tools and treatment techniques. These enhanced theoretical frameworks provide a more logical rationale for evaluating and building upon the results of empirical research results.

5.3. Principle 3: Comorbidity

People are living longer lives, which makes them vulnerable to multiple chronic diseases (comorbidities). Among the ever-increasing number of persons surviving one or multiple strokes, there is a concurrent increased risk because their systems are more vulnerable to developing depression, dementia, and delirium. Recent epidemiological evidence suggests that chronic conditions during mid-life, such as high cholesterol, hypertension, and diabetes, place people at much higher risks for developing dementia in later years. Thus, issues of comorbidity are often central to applied geropsychology and need to receive increasing attention.

See Also the Following Articles

Age-Related Issues among Minority Populations ■ Aging and Competency ■ Aging, Cognition, and Medication Adherence ■ Anxiety Disorders in Late Life ■ Behavioral Medicine Issues in Late Life–Pain, Sleep, etc. ■ Cognitive and Behavioral Interventions for Persons with Dementia ■ Cognitive Skills: Training, Maintenance, and Daily Usage ■ Dementia in Older Adults ■ Depression in Late Life ■ Elder Caregiving ■ End of Life Issues ■ Personality and Emotion in Late Life ■ Psychotherapy in Older Adults

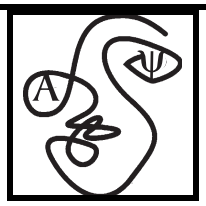
Further Reading

Curry, L., & Jackson, J. (2003). *The science of inclusion*. Washington, D.C.: Gerontological Society of America.

Kane, R. L., & Kane, R. A. (2000). *Assessing older persons*. Oxford, UK: Oxford University Press.

Lichtenberg, P. A. (Ed.) (1999). *Handbook of assessment in clinical gerontology*. New York, NY: Wiley.

Lichtenberg, P. A., Murman, D. L., & Mellow, A. (2003). *Interdisciplinary handbook of dementia: Psychological, psychiatric and neurological perspectives*. New York, NY: John Wiley & Sons.



Gestalt Therapy

Stephen G. Zahm and Eva K. Gold

Pacific University School of Professional Psychology, Forest Grove, Oregon, USA, and Gestalt Therapy Training Center–Northwest, Portland, Oregon, USA

1. Description of Treatment
2. Theoretical Bases
3. Applications and Exclusions
4. Empirical Studies
5. Case Illustration
6. Conclusion
Further Reading

GLOSSARY

background As distinguished from figure or foreground, that which is not uppermost in awareness; provides the context and is the source from which figures emerge and become prominent in awareness.

closure The result of the natural movement toward the completion of perceptual or experiential units.

contact The experience at the self/environment boundary that leads to assimilation and growth. Good quality contact involves awareness and excitement.

contact boundary The dynamic relationship at the meeting point of self/other or self/environment; where experience occurs, and the focus of therapeutic intervention.

creative adjustment A person's best adaptation to dealing with feelings and needs in response to environmental inadequacy in meeting the needs or competing environmental requirements.

figure What stands out from the background as a focus of interest and attention; generally related to a need or desire.

figure formation/destruction process The process whereby a figure of interest emerges, becomes a focus of attention, and is acted upon, resulting in the dissolving of that figure.

foreground As distinguished from background, that aspect of experience that is the focus of attention at any given point.

gestalt A German word that has no direct translation in English. "Configuration," "structure," and "whole" are used as translations, but none of these capture the complete meaning. "Figure" and "gestalt" are often used interchangeably.

organismic self-regulation The process whereby homeostasis is maintained via the figure formation/destruction process.

Gestalt therapy has roots in psychoanalysis, Gestalt psychology, Reichian character analysis, and the work of the early phenomenologists. It was developed in the 1940s primarily by Frederick Perls, Laura Perls, and Paul Goodman. Gestalt therapy is humanistic, holistic, and experiential. It is a process-oriented and relational approach based on a theory of healthy functioning. The focus in therapy is on disruption of the natural process of self-regulation. The goal of treatment is the restoration of awareness, which allows for increased choice and flexibility in all aspects of living.

1. DESCRIPTION OF TREATMENT

1.1. Philosophical Foundations

Gestalt therapy method is guided by three philosophical foundations. The first of these is field theory, a concept that derives from physics. The "field" is a dynamic interrelated system, each part of which influences every

©2004 Elsevier Inc.
All rights reserved.

other part. Nothing exists in isolation. Gestalt therapy is a field theoretical and process-oriented approach, which means the therapist attends to the total field, including content and subject matter, as well as the here-and-now process occurring in the moment. This involves things such as the patient's tone of voice, style of communication and interaction, facial expression, physical gestures, posture, breathing, sensation, and affect. The therapeutic field also includes the patient's history and current situation, the therapeutic relationship, feelings experienced by the patient and the therapist, and their mutual impact on one another.

The second philosophical foundation is phenomenology. The Gestalt therapist attempts to bracket off preconceived biases, beliefs, theories, and interpretations and attend as much as possible to what is actually understandable through the senses, with the focus on the patient's subjective experience. The therapeutic stance is one of meeting the patient where he is, and the therapist describing what she observes, rather than explaining or interpreting. The patient is encouraged to do the same. The therapist also brings her own feelings, sensations, and awareness into the interaction with the patient when useful for the treatment.

The third philosophical foundation is dialogue. A dialogic stance requires that the therapist be "present," confirm the patient, and be available for an I–Thou way of relating. Presence is the therapist bringing all of himself to the encounter in the here and now. Confirmation involves seeing and accepting patients for all they are and all they are capable of being. The I–Thou mode of relating has the qualities of immediacy, directness, presence, and mutuality. These qualities create a therapeutic relationship that is not hierarchical and a meeting between patient and therapist that is in and of itself healing.

1.2. Goal of Treatment/Theory of Change

In Gestalt therapy, the goal is increased awareness. Although the patient presents specific symptoms, such as anxiety, depression, stress-related physical complaints, relationship difficulties, or character/personality issues, the Gestalt therapist does not focus only on symptoms or behavior change. Rather, Gestalt therapy provides a holistic perspective in which symptoms and problematic behaviors are seen as the person's best attempt to deal with conflicting needs or needs conflicting with environmental requirements. Psychological problems, symptoms, and difficulties in interpersonal relationships result from what were originally creative adjustments that have become rigid, reflexive, and unaware, restricting the experience of self and possibilities. By focusing on moment-to-moment process in the therapy session, the Gestalt therapist works with how the person creates and maintains her particular experience of self and other and how this experience impacts the current situation.

Change then occurs spontaneously as a result of this increased awareness. Gestalt therapy's change theory is paradoxical in the sense that change occurs not as a result of attempting to change (one part of the self trying to control and dominate another part) but as a result of the patient attending to, investing more fully in, and living more completely the actual current experience—what is. Doing this enables new responses to emerge that better fit current needs. Increased awareness leads to choice and opens up new options for behavior and interaction.

As aspects of the personality that have been disowned are reidentified with, an expanded sense of self is also restored. For example, the patient increases her capacity to experience and express grief, anger, or tenderness, which expands her ability to live authentically and to connect more with other people. Gestalt therapy does not focus directly on problem solving but rather on what limits the patient's capacity for creative solutions. As the saying goes, "Give a person a fish and he eats for a day; teach a person to fish and he eats for a lifetime." A Gestalt therapist is interested primarily in the latter.

1.3. Therapy Method

By attending to observable behaviors that indicate blocked awareness and interruptions to acting on needs, desires, or interests, Gestalt therapy captures the essence of a person's existential position in the world, which impacts all aspects of self-experience and relationships. For example, the patient speaks softly, apologizes frequently, becomes slightly tearful but cannot cry, changes the subject, or shuts down if strong feelings threaten to emerge. Can the patient become aware of how she clamps down on herself? What makes this shutting down necessary? How might holding back or not allowing herself to take up space be a theme in this patient's life? How many areas of her experience and relationships are impacted by it? How is it related to her presenting symptoms of anxiety or depression or to the dissatisfactions in her life?

The therapy process enables increased awareness of these blocks and how they are maintained and also examines why they were initially developed. For example, if the experience or expression of sadness has been criticized, punished, or even ignored by the parent, the child learns to cut off the feeling, its expression to others, or both. If a child's needs have been met with resentment, criticism, or disgust, the child's creative adjustment is to learn to block awareness of needs or to learn to avoid showing or expressing them to others.

Through focused attention, the patient can become aware of how he blocks feelings/expression, for example, tightening muscles to hold back tears and also internal messages such as "Don't be weak" or "Men don't cry." The patient will discover the etiology—what made this necessary in his development. Symptoms then resolve as the patient reconnects with his authentic self and experience. For example, anxiety may result from a fear that certain "unacceptable" needs or feelings will surface in a particular situation. Depression may result from holding in anger or grief or from poor self-esteem from directing frustration or criticism toward the self that needs to be expressed to another. Relationship problems may result from an inability to open up and reveal feelings or difficulty needing another person.

Gestalt therapy's method is experimental and experiential. Experiments can be diagnostic as well as serving to highlight current experience and uncover new possibilities. For example, an unassuming person who does not make eye contact and qualifies most of what she says complains that people do not listen to her or take her seriously. In therapy she might be asked to experiment with lecturing the therapist on the proper way to do something that is in her area of expertise. She might then show and experience the side of herself that is a knowledgeable expert. She may begin to speak more directly and forcefully. In this process she can get more in touch with this aspect of herself as well as increase her awareness of her reluctance to show this aspect of herself to others or her anxiety in doing so. Conversely, the patient may speak to the therapist as if she is not an expert, when in fact she is. Her language might be filled with reflexive use of qualifying phrases. She might speak without conviction. In either scenario, the patient can see and understand how she contributes to others not giving her the respect she desires.

The patient also learns what made this way of being necessary. She could have been told "Children should be seen and not heard," or one of her parents might have become angry or silently disapproving if she expressed an opinion of her own. She might have been told she was stupid or wrong. Prior to therapy, she may not have remembered these incidents, she may have dismissed them, or she may have minimized their impact. She may not even have been aware that she presented herself as timid and unsure of herself or why it was important to do so. With the new awareness, she regains the possibility of a wider range in experience and behavior.

Gestalt therapy uses experimentation to increase awareness. Meaning for the patient then comes from experienced awareness rather than being imposed by the therapist's interpretation. The therapist tries to have no goal, agenda, or desired outcome in proposing an experiment, other than to observe and work with what happens next. An experiment can also be used to highlight an aspect of the patient's way of communicating. For example, a person who qualifies much of what he says by use of words such as "maybe," "possibly," or "I guess" could experiment with either exaggerating the use of qualifying words or eliminating them entirely and adding after every sentence "and I mean that." The therapist encourages the patient to pay attention to his experience as he does an experiment and to express any reluctance in doing it.

Experiments can also enrich what a person is saying and transform it from "talking about" to a lively present encounter. For example, the patient might imagine the person she is talking about is present in the room and express her feelings directly to that person. This use of the empty chair or two-chair experiment can also be effective in highlighting an internal conflict. The patient might imagine a part of herself in the empty chair and create a dialogue between the conflicting aspects of herself, such as the part that wants to be more creative and take more risks and the part that is more logical and conservative. This type of experiment can also explore an area of impasse or stuckness, for example, suggesting the patient develop a dialogue between the part of herself that wants to leave her marriage and the part that wants to stay.

As an example of some of the previously discussed concepts, a patient who learned to stifle her feelings and needs and to nurture herself with food enters treatment depressed and with a poor self-concept, complaining that she has "tried everything" to control her eating and has been unsuccessful. Turning to food to meet emotional needs was at one time her creative adjustment; it was the best option available given the needs she had and the lack of available gratification

possibilities—in this case, a mother who was stern and unaffectionate and sent the patient to her room if she was "too emotional."

The Gestalt therapist observes this patient's responses in the moment and helps to facilitate her awareness of them. For example, the therapist notices that the patient speaks more quietly and holds her breath when she talks about a hurtful or frustrating experience. The therapist can either make an observation ("It looks like you're tightening your jaw" or "I notice you holding your breath") or ask the patient to attend to what she is currently aware of ("What is your experience right now as you tell me this?"). The patient becomes aware that she is trying to control herself. The therapist attends to her need to control herself. He also attends to what she is controlling—verbalizing feelings and showing anger or tears. Through this process the patient becomes aware that there is also a part of her that would like to express itself and feels held back. Both aspects are explored.

The working-through process involves the patient becoming aware of the importance of holding back feelings to prevent reexperiencing the rejection she felt from her mother or the anger she felt toward her mother when she had feelings or needs. She might remember a forgotten traumatic experience. She might also discover that she expects the therapist to have a similar response. This patient learned over time to identify with her "nonfeeling, nonneeding" self. Her eating behavior may be one of the only remaining vestiges of her "feeling, needing" self. In therapy, she can also experiment with giving voice to the part of her that wants to eat and reidentify with her needs and desires, instead of only wanting to suppress these needs. For example, she might realize that although part of her wants to stop overeating, part of her—out of her awareness—also wants to continue eating as a way to take care of feelings and to comfort herself. This part might say something like "Don't take food away from me, it's all I have." Change becomes possible as she invests more fully in being where she is—feeling those needs she translates into a desire for food—and may take the form of learning to nurture herself in other ways or to get her needs met in relationship with others.

2. THEORETICAL BASES

Many central concepts of Gestalt therapy theory are based on laws of perception discovered and studied by Gestalt psychologists in the early 1900s. Gestalt therapy applies these laws more broadly to all aspects of experience and psychological functioning. Importantly, these Gestalt psychologists discovered that we are not simply passive recipients of perceptual stimuli, but that we are active in organizing our perceptual field. Laws discovered by the Gestalt psychologists include the natural tendency to perceive a figure against a background as a way of organizing experience and a natural tendency toward completion or closure.

Gestalt therapy takes its theory of healthy functioning from a biological concept called organismic self-regulation. This describes the organism's process of taking in from the environment what is needed (food and oxygen) and expelling into the environment that which is not required (waste products and carbon dioxide) in order to maintain balance or homeostasis. Applied to psychological functioning, this theory states that in health, people will naturally go to the environment to get emotional and psychological needs met—the need to be listened to and the need for support, sex, physical comfort, and social contact—and will discharge as needed by talking, expressing feelings, crying, touching, doing, and creating. Gestalt therapy theory does not view the individual as separate from the environment but rather considers the individual/environment field.

According to Gestalt therapy theory, such self-regulation occurs via the figure formation/destruction process. This process is central to any human functioning and as such is a core characteristic of people. In all functioning there is an ongoing process of the formation of a figure of interest and the eventual gratification and dissolution of this need or interest. This cycle has several stages: (1) awareness, (2) clarification of need/interest, (3) scanning self-environment, (4) action, (5) contacting, (6) assimilating, and (7) withdrawing. For example, I am on my way to my favorite restaurant for lunch and see an old friend across the street. I become aware of sensations of excitement and increased energy (stage 1). I recognize my interest in talking to my friend (stage 2). I understand that walking across the street will facilitate meeting this need (stage 3). I walk across the street (stage 4). I say hello and give my friend a hug, and we talk (stage 5). As we say goodbye, I tell my friend and acknowledge to myself what this meeting meant to me (stage 6). I walk back across the street with a feeling of satisfaction, resolve to see my old friend more often, and my focus shifts to what I want to eat for lunch (stage 7).

In healthy functioning, a person will maintain balance by intake and discharge as required for the

satisfaction of emotional and psychological needs by this continuous process of figure formation/destruction. Assuming there is no interference in this process (a disruption that causes the person to misperceive her needs, an environment that is hostile to the needs, or both), what becomes figural or foreground is based on the most pressing need at any given moment. After a need is met, this figure recedes (closure) and is replaced by whatever is next in the hierarchy of needs.

When there is a disjuncture between needs and what is available in (or required by) the environment, a person adapts. A creative adjustment is the best possible accommodation she can make at the time, given her experience, perceptions, and limitations. Over time, these creative adjustments may become obsolete, rigid, and maintained out of awareness, limiting a person's potential for satisfaction and growth. Psychotherapy involves examining and understanding how the natural self-regulatory process has become disrupted, preventing closure on units of experience.

Disruption can occur at any stage of the figure formation/destruction process and diminish a person's experience of himself and his ability for contact. A person who is experiencing disruption interferes with his own development of a strong and clear figure (stages 1 and 2) or with the ability to maintain the excitement and energy necessary for action (stages 3 and 4) or to move into satisfying contact (stage 5), which allows for assimilation (stage 6) and closure (stage 7). Either the interference with the development of a strong figure or the blocking of excitement and action can interfere with good quality contact. A Gestalt therapist is trained to observe both where in the cycle the disruption occurs and how it occurs. Often, the focus is on how the disruption dilutes or distorts interpersonal contact.

In health, a person will perceive, experience, identify with, and act on needs and desires. Healthy functioning requires the ability to readily identify "This is me" or "This is for me" versus "This is not me" or "This is not for me." When disturbances exist in perceiving, experiencing, identifying with, and acting on needs, this loss of functioning is observable by the trained therapist at the contact boundary, which refers to the dynamic relationship at the point where the person and environment meet and interact.

The five processes discussed next are the "how" of any disruption of contact. These processes can occur at any stage of the figure formation/destruction process. It is via these processes, maintained rigidly and out of the person's awareness, that the experience of self and the ability for contact that includes satisfaction of needs and closure is diminished.

Introjecting is experiencing something in the environment as if it is part of the self. An introject is an idea that has been "swallowed whole" without the "chewing" necessary to assimilate it and make it truly a part of the self. Introjecting is particularly problematic if it is a person's primary way of dealing with the world and relationships (such as doing things for others and ignoring one's own needs out of a sense of duty or obligation) or when the introject is destructive (such as a parental introject "You're good for nothing").

Projecting is disowning a feeling, behavior, attitude, or trait of one's own and attributing it to another or the environment. For example, a person who is not able to be angry—due to an introject—may inaccurately perceive another person as angry at him.

Retroflecting is when energy or action that could be directed toward the environment is directed at oneself. Retroflecting is of two types. In the first type, a person directs toward herself a feeling or action meant for someone or something in the environment. For example, the person who is unable to experience anger toward another person—again due to an introject—might end up blaming herself. Any unaware suppression or shutting down is also accomplished by this type of retroflecting. The second type of retroflecting is doing for/to oneself what one would like another person to do for/to oneselft. For example, a person might soothe herself, buy herself gifts, or hug herself instead of getting these needs met by interacting with another.

Confluence is the attempt to deny the existence of a self/environment boundary. Confluence requires a lack of discrimination and articulation of points of difference or otherness. Although there is the possibility of a "healthy" confluence during which two people feel so joined that they seem "as one" (such as at the moment of orgasm) and there is momentary dissolving of the boundary, prolonged attempts to maintain confluence prevent satisfying contact.

Deflecting is behavior that dilutes or reduces the intensity of contact. Examples are avoiding eye contact, laughing off what one says, circumlocution, and understating one's true feelings. There is disagreement among Gestalt therapy theorists as to whether this is to be considered a separate self-regulatory boundary process or whether it is the behavioral manifestation of the others—that is, a result of introjects, retroflection, projection, or desire to maintain confluence.

3. APPLICATIONS AND EXCLUSIONS

Because Gestalt therapy is not an adjustment therapy and strives to be value-free, with a focus on understanding each unique individual and the individual's subjective experience as opposed to an "objective" standard for psychological health, there is no population that would be excluded from treatment. The phenomenological and field theoretical stance, whereby the therapist is striving to understand the patient's subjective experience rather than impose his own, make this orientation valuable for different cultures and minority groups. The Gestalt therapist does not claim to know what is right for the patient, nor does he desire any particular outcome: The only goal is increased awareness for the patient that leads to more choice.

Gestalt therapy can be used with all patient populations (children, adolescent, adult, geriatric, physically/mentally challenged, prison inmates, and inpatient and outpatient psychiatric patients) as well as in a variety of modalities (individual, couples, family, group, and organization). It is applicable to a wide array of problems [e.g., eating disorders, substance abuse, affective disorders, personality disorders, posttraumatic stress disorder (PTSD), adjustment disorders, loss, and grief]. It is an approach that can be used in both long- and short-term therapeutic work. Of course, not every therapist is equally effective with all patient populations, modalities, and diagnoses. The experienced and well-trained Gestalt therapist would be aware of personal limitations and areas of expertise and refer those patients who fall outside of these limits to other practitioners, just as a responsible therapist of any orientation would.

4. EMPIRICAL STUDIES

Research on the effectiveness of Gestalt therapy is scant and primarily confined to specific aspects of the approach. Most research has been focused on one technique that derives from Gestalt therapy method—the empty chair (or two-chair) experiment. It is important to note that use of this or any other single technique does not actually constitute a Gestalt therapy approach, without application of the three philosophical principles described in the treatment section. Although no single technique can be said to represent the essence of the Gestalt therapy approach, a technique has the advantage of being easily defined and thus accessible to the manipulations of empirical research.

The majority of research used the empty chair technique. Subjects were asked to have a dialogue with either a part of themselves or someone else relevant to the issues with which they were grappling. Most of the studies that used the empty chair technique found it to be more effective than other forms of therapy or control groups. It was also found to be more effective in reducing anger and producing positive attitudes than either intellectual analysis or emotional discharge. It was more effective than empathic reflection or focusing for creating a greater depth of experience. For resolving "unfinished business," it was more effective than an attention-placebo condition.

In one study, K. M. Clark and L. S. Greenberg found the empty chair technique to be more effective than either a cognitive decision-making task or a waiting list control group for resolving decisional conflict. During the process of using this technique, the original decision the person had wished to make often unfolded into a deeper related decision.

In summary, studies have shown this technique to result in an increase in affect, depth of experience, and the resolution of related emotional issues. Results have included a decreased sense of decisional conflict, reduced anger, increased positive attitude, lower systolic blood pressure, and the resolution of unfinished business.

Two studies found no benefit in using the empty chair technique for resolving grief issues or dealing with mild depression. In a study conducted by N. P. Field and M. J. Horowitz, the intervention consisted of placing subjects in a room alone and asking them to talk to their deceased spouses for 5 minutes through the direction of a taped recording. Nothing like this would ever be done in an actual Gestalt psychotherapy, which requires a focus on moment-to-moment process and the presence of the therapist. This technique was not found to be effective in reducing symptomatology 6 months later. Another study found the empty chair technique to be no more effective than time for decreasing mild depression. The researchers stated that the lack of positive results may have been due to the short duration of treatment, gravitation toward the mean over time, or experimenter bias.

Focused expressive therapy (FEP) was used in several research studies. FEP borrows from Gestalt therapy method and uses such procedures as directed fantasy, two-chair dialogues, and awareness exercises. It emphasizes the importance of emotional insight and the magnification of internal experiences, but it also includes a high level of authoritative guidance and

confrontation that is inconsistent with a Gestalt therapy approach. FEP was not found to be as effective as either cognitive group therapy or self-directed therapy. However, the description given for self-directed therapy was actually more similar to Gestalt therapy as it is practiced than the more authoritarian FEP method used to represent Gestalt therapy in this study.

Few studies have examined a more complete Gestalt therapy approach. Most of these yielded positive results, such as an increase in levels of self-actualization and more positive attitudes toward body image. Gestalt therapy was found to increase internal locus of control and the assumption of personal responsibility in prison inmates. A Gestalt therapy approach resulted in more behavior change than empathic reflection. Gestalt therapy was found to help Vietnam veterans with symptoms of PTSD. In couples therapy, it was effective in replacing power struggles with self-definition, limit setting, and intimacy, resulting in greater clarity, acceptance, and understanding.

5. CASE ILLUSTRATION

The following case example illustrates the treatment process and application of the theoretical concepts that have been presented. The therapist in this case was S. Zahm.

5.1. Kim

Kim, a 44-year-old successful professional woman, was referred by her physician due to recurrent periods of depression. She had been depressed off and on for much of her life. The bouts of depression had worsened recently in both duration and severity. She had not been able to tolerate antidepressant medications, and two previous experiences with therapy, both brief, had not yielded lasting improvement.

Kim presented as a smart, attractive, stylishly dressed woman. Her wit and humor were evident despite her somewhat depressed affect. Her facility with language, keen observations, and easy manner made it clear why she had achieved success professionally. During the initial session, Kim asked perceptive questions about my experience, my therapy orientation, and how I thought therapy could help her. She had been disappointed in her previous therapy experiences. When asked what she hoped to get from therapy, she answered "Not just a Band-Aid. I'd like to get to the bottom of what this depression is about."

Kim described herself as "an overeducated overachiever" who had been successful in school and in her profession. She worked long hours but wondered what was the point of all her hard work because she took little enjoyment from her accomplishments. Kim saw her life as lacking in meaning, with little that gave her pleasure. She said she felt "down" a lot, woke up in the morning dreading the day ahead, and said she was just "going through the motions" of social interaction.

Kim described her relationship with her husband Bill as "pleasant enough" but lacking a sense of emotional connection or passion. She reported little interest in sex and stated that their infrequent lovemaking was "more for him." She said "it's like I'm dead inside, or just numb." Kim reported that most of her friends were work acquaintances, and that she had only one "good friend" but did not really confide in her. Kim and her husband had no children. Kim and I agreed to meet for weekly individual psychotherapy sessions and to track whether she felt the sessions were useful and whether she was getting what she needed.

In describing her history, Kim reported that her parents had divorced when she was 8 years old. Her father abandoned the family, moving to another state to avoid paying child support. Kim, the oldest of three, became responsible for helping her mother with her younger siblings and doing household chores so that her mother could go back to work to support the family. Kim also became her mother's confidant and emotional support. She reported that she never really "got to be a kid" after that, and that her mother "never had time" for her. Kim's mother died approximately 8 months prior to our first meeting. She minimized her feelings about her mother's death, saying she thought she had "dealt with it."

Kim's therapy themes emerged right away in her relationship with me. In our initial sessions, her fear of showing or revealing vulnerable feelings to me stood out. For example, when Kim talked about her mother's death she would reflexively shut down. I noticed her retroflections and deflections, including choking back tears, changing the topic, and looking away. I pointed these out in an attempt to sharpen this figure, to clarify what was most important to her at that point. I observed that although she had strong feelings about this loss, it was also important to her not to show these feelings or talk about them to me right then.

This process helped Kim focus on her reluctance to open up and share vulnerable feelings. The theme of having a variety of "softer" feelings (sadness, loss, and emptiness) and the reluctance to express them to me was something we worked with over the course of the

therapy. The essential components involved introjects such as "Be strong" and "Don't show weakness" and the belief that no one would be interested in her feelings.

Within the first few months of therapy, Kim became more aware of these internal mandates and saw how they were connected to her relationship with her mother. Kim had learned not to reveal her feelings and emotional needs but rather to focus on taking care of her mother and siblings and deal with any of her needs by herself. Initially, it was a major step for Kim to identify that she felt a need and to also feel her difficulty and discomfort in expressing it to me.

I suggested Kim experiment with looking at me and imagining allowing herself to let down and cry with me. She imagined me having a critical and rejecting response. As she worked with her reluctance over a period of weeks, and felt supported by me in doing only what she was comfortable with, her reluctance moved into the background and her organismic desire and need to experience and express her feelings became more figural. At that point she was able to let down into her feelings of loss and grief, thus beginning the process of getting the closure she needed with regard to her mother's death.

Afterwards, Kim was surprised to discover that this felt OK. It was still difficult for her to ask me how I now felt toward her. Her strong pull was to be satisfied with not knowing how I felt and what my reaction had been to her crying. I pointed out how not asking me about that left her with a blank space that might then be filled with her own fears. She decided to ask the question, "How do you feel about me now?" I was touched by her courage in asking the question and told her so, and I gave an honest response to her question. This was a pivotal moment in the therapy because Kim could take in and assimilate my genuine caring for her.

As we did this work and Kim opened up to deeper levels of feeling, another theme became apparent to both of us. Kim became fearful of relying on me and afraid that I would leave her if she needed me. She had a first-class radar system that detected the slightest sign of my lack of interest, lack of energy, or attention wandering. For example, if I looked at the clock she became silent and withdrew. When I told her of planned vacations, her response was to joke that she wondered if I would want to come back. When we explored this, it became clear that my going away evoked fear that I would abandon her or lose interest in her.

We worked with this theme in a number of ways. As Kim focused on feelings and sensations, she was able to remember particular events with her father, such as his leaving the house in a flurry of loud words and emotion and yelling things such as "I never get any peace in this house!" One time, Kim tried holding on to him to keep him from leaving and he pushed her away, saying something like "Why can't you kids leave me alone?" The meaning Kim extracted from these experiences and his ultimate abandonment of the family was that it was her fault that her father had left and that she must have been unlovable and "too much" for him.

As we worked with these awarenesses, Kim could feel a little girl part of her that "had always been there" that she tried to ignore. I asked Kim to focus on this sense of herself and to talk to me as the little girl. What emerged from this was that she felt shameful and desperately wanted to hide or disappear. She went back to her adult perspective and told me she wished that little girl would disappear or at least "grow up." I suggested a dialogue between the adult Kim and the little girl. As the dialogue developed, the adult Kim began to feel sad and realized that rather than wanting the little girl to disappear, she wanted to hold and comfort her. She did not see her as "unlovable" or "too much" but instead felt compassion for her loss and pain and was able to tell her so. This newfound ability to experience compassion toward herself was a key in helping Kim integrate her softer "little girl" feelings with her "adult." This occurred over a number of months as she continued to work with these aspects of herself.

When we worked directly with Kim's father she was able to, in fantasy, tell him she was hurt and angry that he had not loved her enough to stay in her life even if he had to leave her mother. She finished her tearful dialogue with him by saying, "It wasn't my fault you left. You failed me, not the other way around." She also became aware that she had never been able to grieve the loss of her father and that her mother would not tolerate her tears and told her she had to be strong and that crying "wouldn't do any good." At that point, she was able to grieve the loss of her father with me and have more of a sense of closure.

In our sessions Kim came to see how the belief "I'm too much" affected major aspects of her emotional and interpersonal existence. This was especially evident in how she experienced her relationships with men and the type of contact she could have with them. She saw that with her husband, she made no demands, rarely got angry, and focused her attention on his needs to the exclusion of hers. Our work

included Kim experimenting with new ways of interacting with her husband, such as being more assertive with her needs. Initially, this made her anxious. As we worked with the anxiety and she realized the responses she feared were not likely ways that her husband would respond to her, she became less anxious and more excited and playful with the possibilities of being bolder about asking for things she wanted. This affected all aspects of their relationship, including their sexual relationship.

Slightly more than 1 year into the treatment, Kim reported that she "couldn't remember the last time" she had felt depressed. She said she felt better about herself, her relationships, and the possibilities for her life than she ever had. She began to spend less time working and signed up for a beginning ballet class, which was something she had "always wanted to do." She developed closer relationships with a few women friends, revealing more personal things and accepting support from them.

I asked Kim to reflect on where she was now and where she had been when she entered treatment. Kim reported that she felt more at peace with her parents, even though she still felt sad about her childhood. She said that it had been "life changing" to risk opening up to another person and allowing herself to need and depend on me. Kim saw the connection between her process of shutting down on her feelings—her self-criticalness, rejecting the little girl part of herself, avoiding feeling her anger and grief, and not allowing herself to need anything from others—and the depression she had experienced for so long.

Kim and I agreed that she was ready to terminate therapy and we allowed 1 month for this process. Given Kim's issues of abandonment and loss, it was important to set an ending date and not dilute the fact that we were ending by meeting less frequently. In this way, Kim had sufficient support and a structure for dealing directly with the loss of her relationship with me. We celebrated her gains together and cried together. At our last session, we said good-bye in a complete way that Kim said she had never had the opportunity to do with anyone else.

5.2. Outcome

Growing up, Kim experienced a lack of awareness of and responsiveness to her emotional needs from both of her parents. Her father abandoned Kim completely, and her mother was unavailable for Kim's needs, requiring Kim to attend to her needs instead. Kim's creative adjustments involved both the awareness stages and the action/expression stages of the figure formation/destruction process. She learned to cut off the experience of certain feelings and to avoid showing and expressing them. The meaning she made of her father's leaving ("I'm unlovable" and "I'm too much") and introjects such as "Be strong" created her retroflective process of cutting off and alienating aspects of herself, focusing on the needs of others, and not showing or expressing her own feelings and needs. This process resulted in unfinished emotional situations in which she had not been able to get closure. It resulted in the sense of emptiness, lack of meaning, and numbness that she described at the beginning of treatment as well as in unsatisfying interpersonal relationships that lacked depth. It also contributed to her career success, which was in part based on her skill in attending to the needs of others and her willingness to work hard so that she would not be rejected or abandoned.

All these issues emerged and were worked on in the context of the therapy relationship. As Kim's awareness increased, she was able to work through her fears and reluctance and to feel the need to express and show her authentic feelings. She came to feel compassion for a part of herself she had previously cut off and wanted to be rid of, and through this experience of compassion she was able to reidentify with and integrate this part of her that included needs and softer feelings. She was able to redirect her self-criticism and see that it was not her being "too much" that caused her father to abandon her and her mother to be unable to attend to her needs. She came to feel that her little girl had deserved to be better taken care of.

In Kim's relationship with me, she could experience my genuine caring and also see that when I was not there for her (e.g., when I went on vacation) it was not because she was unlovable or "too much." She learned to express disappointment with my limitations, or anger at me for leaving, in a way that she had never had the opportunity to do with her parents.

The therapy carried over to and influenced every aspect of Kim's life and interactions. Her relationship with her husband deepened, and she began to consider what she wanted to do in her life besides working. However, it was Kim's relationship with herself that changed most dramatically. Kim no longer criticized herself or felt she was weak if she had feelings and needs. She no longer cut off and numbed a major part of who she was. She had greater range of emotional experience and expression. Kim understood that her

lifelong depression was the result of ignoring, denying, and cutting off her feelings, which led to an inability to identify and act on her needs. This resulted in a lack of closure on difficult emotional situations as well as unresolved issues of grief and loss.

6. CONCLUSION

The practice of Gestalt therapy is based on the philosophical foundations of field theory, phenomenology, and dialogue. Gestalt therapy is a process-oriented approach that focuses on the person/environment field. The goal in Gestalt therapy is increased awareness, and change occurs through focused attention on "what is" and not by attempting to achieve a particular goal or agenda. Because the therapist brings himself or herself to the meeting with the patient, a type of contact that is in itself healing becomes possible. Gestalt therapy is based on a theory of health, and the focus is on how the patient interrupts his or her natural self-regulation process. The method is experiential and experimental, relying on meaning emerging from experience rather than interpretation. Gestalt therapy has wide applicability with a variety of patient populations, problems, and treatment modalities. Although limited, research indicates its usefulness for a variety of populations and issues.

See Also the Following Articles

Ideological Orientation and Values ■ Mental Measurement and Culture ■ Stress ■ Values and Culture

Further Reading

Beisser, A. (1970). The paradoxical theory of change. In J. Fagan, & I. Shepherd (Eds.), *Gestalt therapy now*. Palo Alto, CA: Science and Behavior Books.

Breshgold, E. (1989). Resistance in gestalt therapy: An historical/theoretical perspective. *Gestalt Journal*, *12*.

Jacobs, L. (1989). Dialogue in gestalt theory and therapy. *Gestalt Journal*, *12*.

Kepner, J. (1993). *Body process*. San Francisco: Jossey-Bass.

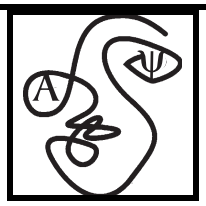
Latner, J. (1988). *The gestalt therapy book*. Highland, NY: Gestalt Journal Press.

Polster, E., & Polster, M. (1993). *Gestalt therapy integrated*. New York: Brunner/Mazel.

Wheeler, G. (1991). *Gestalt reconsidered*. New York: Gardner Press.

Yontef, G. (1993). *Awareness dialogue and process: Essays on gestalt therapy*. Highland, NY: Gestalt Journal Press.

Zinker, J. (1997). *Creative process in gestalt therapy*. New York: Brunner/Mazel.



Gifted Students

John F. Feldhusen
Purdue University, West Lafayette, Indiana, USA

1. Introduction
2. Defining Gifted
3. Why Separate Students?
4. Identification and Nurturing
5. Worldwide Support
6. Literature Regarding Gifted Students
7. Current Research
Further Reading

GLOSSARY

aptitude Specific ability in an academic area such as mathematics, spatial tasks, science, or verbal activities.

elitism An attitude critical of special programs for gifted and talented youth because they provide benefits not available to all students.

enrichment programs Special services for gifted youth that avoid the regular school program and curricula and that often offer projects and creative activities. Students are often pulled out of regular classes during the school day for these programs.

gifted Youth who have high general mental or artistic abilities.

intelligence General mental ability as often represented by gifted students.

live-in experience A service for gifted, talented, and/or precocious youth in which they leave their homes and go to live at the site where a special program is offered, often a college or university campus.

precocity Achievement or performance that surpasses the average performance of youth at a given age level. The term precocious is often used as a synonym of gifted.

talent Specific ability in an applied area such as mathematics, science, leadership, music, art, or writing.

Gifted, talented, and precocious students can be identified with intelligence and aptitude tests, rating scales, achievement test scores, and teacher observations. Once identified, schools can offer special classes, programs, and services to facilitate the growth of their gifts and talents. Colleges and universities also provide special programs and services in the forms of early admission to highly precocious youth as well as special weekend and summer programs.

1. INTRODUCTION

Gifted, talented, precocious, high aptitude, high achieving, and creative are all terms used to designate children and youth of high or superior ability. There is little agreement about test instruments, rating scales, and observational procedures used to identify these children, or cutting levels on the various test instruments that are used. However, there is some agreement that test scores 2 standard deviation above the mean or percentile scores at or above 90% are appropriate identification criteria. There is also agreement that it is desirable to gather two or more measures to identify candidates for special programs.

©2004 Elsevier Inc.
All rights reserved.

2. DEFINING GIFTED

The field of gifted education has generated a large research base, beginning with the work of Lewis Terman. Shortly after his development and publication in 1916 of the first intelligence test in the United States, the Stanford–Binet, he began a major longitudinal study of 12-year-old children with IQs of 135 or higher who were followed throughout their lives, with special interest in their long-term achievements. Another pioneer and researcher in the field was Leta Hollingworth, whose work focused on highly gifted youth and their education. Terman's research established the initial focus on high IQ as the major means of identification, and his following research established that those who go on to high-level creative production in life profited from early grade advancement or acceleration and motivation to achieve at a high level. Hollingworth's work established the critical need of highly gifted youth for challenging and motivating instruction and the problems they face in regular grade-level classrooms, often with teachers who do not understand them and their special needs.

3. WHY SEPARATE STUDENTS?

The major purpose of identification of children for gifted programs or educational services is, or should be, to determine if they need a special regimen beyond the services of a regular classroom program and what particular services or resources might be needed. Decisions about a child's special education needs can best be made through the assessment of the child's special aptitudes and talents or areas of high achievement. Then, if both aptitude and achievement tests reveal high mathematical ability and the need for higher level mathematics instruction, a decision may be made to place the child in the next higher grade level for mathematics instruction or in a special mathematics enrichment program. Alternately, if the child's achievement test scores are all above the 95th percentile, after conferences with the child and his or her parents, a decision may be made to move the child to the next higher grade level for all instruction.

In some schools or communities, gifted or precocious children are grouped together in special classes or even in special schools for gifted children only. Teachers of such classes often have received special training and certification for work with gifted children and youth and use curricula designed especially for gifted and precocious children.

Several states have established schools for precocious secondary students, often with major academic focus on mathematics and science and offering a full-time experience. These schools are comparable to the special day schools offered in several cities such as New York and Chicago, again stressing mathematics and science. Admission may be strongly competitive, but the instruction and curriculum as well as association with other academically superior students make these schools ideal programs for academically precocious youth.

4. IDENTIFICATION AND NURTURING

Much of the attention given to gifted and talented youth has turned diagnostically, in the past decades, to identification and nurturing of their special talents. Rather than offering a general, homogeneous regimen and curriculum for all gifted youth, the new programs offer a variety of classes and services focused on their special and emerging talents. Thus, a child who has been identified as having exceptional art talent can participate in an after-school arts class, another who shows strong leadership ability may enroll in Saturday morning leadership classes, and yet another who is precocious in science can be matched up with a mentor-professor at a community college and spend one afternoon a week at the college working with the mentor-professor in a science lab. Talent-oriented programs use a wide variety of services and resources in the school and community to foster the development of children's special talents.

Saturday and summer programs are also now widely available to gifted and talented children with or without school identification or endorsement. Although some of these programs make no pretense of identifying and serving the gifted, many of them explicitly aim to identify and serve gifted and talented youth. These programs often offer talent-oriented classes taught by experts in that field. For example, an actor or director may teach a class in dramatics, a computer professional may teach a class on computers, a math teacher may teach an advanced math class, an attorney may teach a class on the law as a profession, and a local novelist may teach a creative writing class. These classes supply a knowledge base in the field of the talent, offer a motivating challenge for talented youth, and provide mentors and models

in areas related to youth talents. Some programs, such as one at Purdue University, also offer mentoring experiences as a follow-up to the classroom in areas of youth talent.

5. WORLDWIDE SUPPORT

The field of gifted and talented education is now fostering identification and special education for gifted and talented youth in countries throughout the world through the auspices of the World Council for Gifted Education. Its major publication, *The International Handbook of Giftedness and Talent* (2000), is an excellent resource, with chapters written by the leading experts in gifted education from throughout the world. In the United States, the major professional organization with focus on gifted and talented youth is the National Association for Gifted Children, based in Washington, D.C. Most of the U.S. states also have professional organizations focused on gifted and talented youth.

6. LITERATURE REGARDING GIFTED STUDENTS

There are a number of publications devoted to the nature and educational needs of the gifted and their teachers, notably *The Gifted Child Quarterly*, the major publication of the National Association of the Gifted Children, Washington, D.C.; the *Gifted Child Today*, Prufrock Press, Waco, Texas; the *Journal of Secondary Gifted Education*, Prufrock Press, Waco, Texas; *Roeper Review*, the Roeper Institute, Bloomfield Hills, Michigan; the *Journal for the Education of the Gifted*, the Association for the Gifted, of the Council for Exceptional Children, Reston, Virginia; *High Ability Studies*, the journal of the European Council for High Ability, Bonn, Germany; *Gifted and Talented International*, World Council for Gifted and Talented Children, College of William and Mary, Williamsburg, Virginia; and *Gifted Education International*, AB Academic Publishers, Oxon, United Kingdom.

7. CURRENT RESEARCH

The field of gifted and talented education is diverse in the nature of its research and the direction of educational services. Research runs the gamut from studies of the nature of talent and abilities to personality and attitudinal aspects of giftedness, methods of testing and identifying gifted youth and the appropriate test instruments, creativity and problem solving abilities of gifted and talented youth, gender differences in abilities, models' and mentors' roles in the development of gifts and talents, and nature versus nurture in the development of giftedness. Research on gifted and talented youth and in the field of gifted education is widely reported in the journals of the field and at conferences.

The field of gifted and talented education is also diverse in its conceptions and identification of giftedness, and in its educational services and programs for gifted and talented youth. It is sometimes viewed as elitist by mainstream education, and there is some opposition to special classes and services for these youth. Funding for research in gifted education is limited, but several centers at universities carry on excellent programmatic investigations, notably at the University of Connecticut, the University of Iowa, and Purdue University. It appears that educational programs for gifted and talented youth are now well established in many cities, U.S. states, and other countries, and they will continue to grow and provide valuable educational experiences for gifted youth.

See Also the Following Articles

Intelligence and Culture ■ Intelligence Assessment ■ Intelligence in Humans

Further Reading

Ellingson, M. K., Haeger, W. W., & Feldhusen, J. F. (1986). The Purdue mentor program: A university based mentorship experience for gifted children. *The Gifted Child Today*, *9*(2), 2–5.

Feldhusen, J. F. (1991). Full-time classes for gifted youth. *Gifted Child Today*, *14*(5), 10–13.

Feldhusen, J. F. (1995). Talent development vs. gifted education. *The Educational Forum*, *59*(4), 346–349.

Feldhusen, J. F. (2002). Guidelines for grade advancement of precocious children. *Roeper Review*, *24*, 169–171.

Feldhusen, J. F., & Koopsman-Dayton, D. D. (1987). Meeting special needs of the gifted through Saturday programs: An evaluation study. *Gifted International*, *4*(2), 89–101.

Feldhusen, J. F., Rebhorn, L. S., & Sayler, M. F. (1998). Appropriate design for summer gifted programs. *Gifted Child Today*, *11*(4), 2–5.

Feldhusen, J. F., & Sayler, M. F. (1990). Special classes for academically gifted youth. *Roeper Review*, *12*(4), 244–249.

Gagné, F. (1999). *Tracking talents: Identifying multiple talents through peer, teacher, and self nomination*. Waco, TX: Prufrock Press.

Hansen, J. B., & Feldhusen, J. F. (1994). A comparison of trained and untrained teachers of gifted students. *Gifted Child Quarterly*, *38*(3), 115–123.

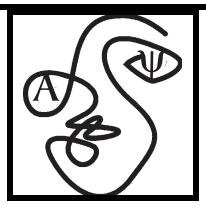
Holahan, C. K., & Sears, R. R. (1995). *Genetic studies of genius: The gifted group in later maturity* (Vol. 6). Stanford, CA: Stanford University Press.

Hollingworth, L. S. (1926). *Gifted children: Their nature and nurture*. New York: Macmillan.

Terman, L. M. (1916). *Measurement of intelligence: An explanation of and a complete guide for the use of the Stanford revision and extension of the Binet-Simon intelligence scale*. Boston, MA: Houghton Mifflin.

Terman, L. M. (1925). *Genetic studies of genius: Mental and physical traits of a thousand gifted children,* Vol. 1. Stanford, CA: Stanford University Press.

Van Tassel-Baska, J. (1992). *Planning effective curriculum for gifted learners*. Denver, CO: Love Publications.



Goal Setting and Achievement Motivation in Sport

Joan L. Duda

University of Birmingham, Birmingham, United Kingdom

1. Goal Setting in Sport
2. The Meaning of Goals
3. Additional Considerations and Future Directions
   Further Reading

GLOSSARY

discrete goal A particular performance standard or objective.
ego-involved goal A goal of revealing superior competence relative to others.
failure-oriented goal-setting style Individual difference reflecting a concern about not demonstrating low ability compared to others and thus a tendency to set very easy or extremely difficult goals.
goal orientation Dispositional tendency regarding achievement goal focus.
goal proximity Having both short-term and long-term goals.
goal-setting style An individual difference factor capturing an athlete's preferences regarding the type of discrete goals that are set in training and/or competition.
moderately challenging goals Standards that are of an optimal difficulty (i.e., not too easy and not beyond one's capabilities).
motor endurance task A physical task that requires muscular or cardiovascular endurance for successful performance.
outcome goal A goal based on a competitive or comparative performance standard.
performance goal A goal to secure some discrete standard or end product that is self-referenced (e.g., a "personal best" achievement).
performance-oriented goal-setting style Individual difference reflecting a tendency to set goals related to learning, performance improvement, and increasing one's self-referent ability.
pre-competitive anxiety The cognitive (e.g., degree of worry) and somatic (e.g., heart racing) anxiety experienced by an athlete before a competitive event.
process goal A goal of trying to meet some standard of proficiency concerning technique or employment of strategy.
success-oriented goal-setting style Individual difference reflecting a tendency to set competitive outcome goals or strive for standards that entail social comparison.
task-involved goal A goal that entails an emphasis on demonstrating personal competence in a self-referenced way and that centers on energetic attempts at task mastery and/or improvement.

Any perusal of applied sport psychology texts, examination of prevailing research topics in major sport psychology journals, or conversational engagement with those involved in sport usually points to the relevance of goals, goal setting, and motivation in this domain. Indeed, investigations of youth, intercollegiate and Olympic athletes, and coaches have supported their propensity for the setting of goals in competition and training. Interestingly, though, this descriptive research also

©2004 Elsevier Inc.
All rights reserved.

indicates that athletes and coaches consider their goal-setting efforts as only marginally effective. That is, although the formulation of and striving for some target(s) or objective(s) appear to be endemic in the athletic context, the employment of goal setting in sport seems less productive than one would readily assume. A major premise of the present contribution is that the implementation of goal setting will remain inconsistent and the efficacy of such strategies unconvincing and restricted in scope unless the meaning of the goals and their link to variations in motivation processes are considered.

In this article, the underpinnings, directions, and major findings of research on goal setting in sport will be summarized. Limitations of this work will also be highlighted. We then move into a discussion of the concept of achievement goals. In particular, attention will turn to the ways in which the constructs and assumptions embedded in con-temporary achievement goal frameworks provide further insight into the types of goals that individuals strive for within the athletic milieu. Moreover, the implications of these different goal perspectives for athletes' motivation to achieve and the overall quality of their sport engagement will be addressed. In essence, it is argued that a consideration of variability in achievement goals helps complete the picture in terms of when goal setting will be effective and what contributes to motivationally sound and adaptive goal setting. The article concludes with a discussion of future directions and considerations regarding work linking goal setting and achievement goals.

1. GOAL SETTING IN SPORT

In the view of Locke and colleagues (1981, p. 126), goals capture what "an individual is trying to accomplish; it is the object or aim of an action." Weinberg (1992) takes this definition a step further by proposing that meeting a goal means "attaining a standard of proficiency on a task, usually within a specified time limit" (p. 178). According to Locke and Latham, we typically are consciously aware of our goals, but, on some occasions, goals can reflect unconscious intentions or purposes. Particularly in terms of the former, it is held that goals initiate and regulate behavior. That is, goals are conceptualized in Locke's model to be the antecedent that is most proximal to performance.

Locke and Latham further argue that goals can have a direct and indirect effect on performance. It is assumed that goals directly influence performance outcomes by centering our attention on task-relevant cues (and attenuating our attending to cues that are not task-relevant) and encouraging persistence and appropriate levels of exerted effort. With respect to the proposed indirect consequences of goal setting, resulting strategic planning and problem solving are held to be mediators of the goal setting and performance relationship. Clearly, the latter mechanisms are especially essential when individuals are trying to meet complicated, challenging, and long-term objectives. However, sometimes focusing on a task, trying one's hardest, and persevering are not enough to secure goal attainment. Before any of these direct or indirect influences are likely to come into fruition, it is suggested that individuals must be committed to the goal(s) at hand and possess the necessary ability or skills. Feedback in relation to goal progress is another proposed moderator that is viewed as being essential to goal-setting effectiveness.

Almost exclusively, research on goal setting in sport has been grounded in Locke's goal-setting theory. The major assumptions of this theoretical framework, particularly in terms of the existent sport-related work, are described next. All in all, these predictions indicate which type of goal is most likely to have the most substantial positive impact on performance.

1.1. Basic Tenets of Locke's Goal-Setting Theory

There are three central propositions in Locke's model of goal setting as it has been tested in sport; namely, that (1) hard goals result in higher levels of performance than do less-demanding goals, (2) specific goals give rise to superior performance when contrasted with no goals, very general goals, or a vague goal such as "do your best," and (3) using short-term goals in conjunction with long-term goals will lead to better performance than using long-term goals exclusively. Overall in the goal-setting literature, and specifically in terms of research conducted within organizational settings, evidence has been garnered indicating that difficult and specific short- and long-term standards result in superior performance when contrasted with more vague goals or no goals. Indeed, a number of large-scale narrative reviews and meta-analyses have heralded the effectiveness of explicit, difficult, short- and long-term goal setting. What is especially impressive about this body of work is that

the value of specific and challenging goals has been supported in the field and in the laboratory, with respect to diverse settings, across different tasks, in the case of various populations, and in terms of short- and long-term performance indicators. Moreover, the proposed direct and indirect mechanisms by which goals are held to influence performance have been generally sustained in the general and organizational psychology literatures.

1.2. Testing of Locke's Theory in Sport/The Physical Domain

In an attempt to encourage more systematic investigation of the efficacy of goal setting in the sport realm, Locke and Latham proffered that the impact of goal setting on sport performance would be even more striking than what is observed in work-related or organizational situations. Further, they proposed a number of hypotheses regarding when, where, and why goal setting functions in the physical domain. With respect to research in real-life sport settings or laboratory experiments involving sport or physical-based tasks, however, the literature on the effectiveness of goal setting is a bit more conflicting and slightly less compelling than has been the case in work settings. In the available sport-focused narrative reviews, it is typically concluded that research has provided only qualified support for the specific standard versus "do your best" hypothesis in Locke's framework, and, on the whole, the findings in terms of goal proximity (i.e., having short-term along with long-term goals) and goal difficulty (i.e., more difficult goals result in better performance) have been equivocal.

In 1995, Kyllo and Landers published a meta-analysis that provided a slightly more optimistic but still rather constrained evaluation of studies on the effect of goal setting on physical task performance. Overall, in terms of the observed relationship between goal setting and performance witnessed in 36 studies, a mean effect size of .34 was revealed. This result is considerably more modest than what has been reported in omnibus meta-analyses not specific to the physical domain (i.e., the latter have found overall effect sizes ranging from .42 to .80). Providing only partial support for one of Locke's propositions, Kyllo and Landers also found moderately challenging goals to be associated with a greater effect size (.53) than difficult (.09) or easy (.07) goals. Consonant with Locke and Latham's hypothesis, combined short- and long-term goals were coupled with a larger effect size (.48) than short-term (.38) and especially long-term (.19) goals only.

In their more recent appraisal of 56 published goal-setting studies in sport and physical activity, Burton *et al.* reported a 78.6% effectiveness rate. As this rate was greater than what was reported in an earlier version of the review chapter in question (in particular, 7 years previously, 66% of the published studies reviewed revealed a positive overall goal-setting effect), Burton and colleagues suggested that more support for the consistency of goal-setting effects is accruing as more research has been undertaken. This is a promising conclusion, but caution is warranted when such a deduction comes from a sample of published papers. That is, published research is more likely to show significant and expected effects when contrasted with non-published studies. Further, it is important to keep in mind that the conclusions of Burton and associates did not stem from a systematic, meta-analytic review.

There has been diversity in the methods used to test the veracity of Locke's theory when applied to sport. Particularly during the earlier work on this topic, more laboratory-based experiments flourished. The trend for such investigations was to focus on a simple, motor endurance-type task (e.g., sit-ups), recruit physical education or sport science majors as subjects, have the experimenter assign the goal(s), and then contrast a specific goal versus a general and/or vague "do your best" goal with respect to the impact on performance, often over a short time period. Sometimes a control group that was assigned "no goals" was incorporated in the design. Other investigations in the sport milieu have been field-based studies and/or interventions employing single-subject designs to examine the efficacy of goal-setting programs.

Another contemporary extension of work on goal setting that emanated particularly within the sport and exercise psychology literature is the distinction between outcome, performance, and process goals. When focused on a process goal, the athlete is trying to meet some standard of proficiency concerning her or his technique or employment of strategy. For a performance goal, the athlete is reaching to secure some discrete standard or end product that is self-referenced (e.g., trying to better one's sprint time). An outcome goal is based on a competitive or comparative performance standard, such as winning a contest. On the whole, sport research suggests that athletes focus on all three types of goals, and they may serve different purposes in terms of maximizing skill development and performance.

1.3. Limitations of Sport Goal-Setting Work

When trying to explicate the mixed results and typically more moderate effects when goal setting is tested in regard to sport or physical task performance, researchers have offered a number of methodological artifacts and design limitations for consideration. These factors, which to some degree can be considered interrelated, include: (1) the inclusion of individuals who tend to be highly motivated in the physical domain as study participants, (2) the tendency for study participants in the control or "do your best" groups to spontaneously set goals, (3) the likelihood of competition being manifested between study participants, and (4) the failure to consider study participants' personal goals with respect to task performance. Without a doubt, such methodological shortcomings pose a serious threat to the internal validity of this research. Moreover, not systematically taking account of such factors in goal-setting research diminishes our understanding of why, when, and how goal setting may work in sport.

Characteristics of the tasks used in sport goal-setting research have also been pointed to as explanations for the compromised effectiveness of goal setting in this setting. Specifically, it has been suggested that physical tasks in and of themselves tend to provide feedback. An individual can see and feel how one is doing and how hard one is trying (especially in the case of those sit-up studies!), regardless of whether one is focused on a more-or-less specific or a more-or-less difficult goal. Both the process of performing and performance outcomes per se are more overt in the physical domain when contrasted to tasks in other settings, such as industrial or academic environments.

According to Hall and Kerr, the anomalies in the sport goal-setting literature hold ramifications for the employment of goal setting in applied sport situations. Addressing the methodological and task-related problems identified above in subsequent research may indeed enhance the internal validity of future goal-setting studies. What would be the cost, however, in terms of external validity? That is, if we controlled for spontaneous goal setting and personal objectives, competition, the probability that those involved in the usually voluntary activity of sport are often motivated to participate and perform, and the nature of the task, what does such research have to say to the real world of athletic engagement?

A central conceptual ambiguity that plagues status quo sport goal-setting work based on Locke's model is the premise that goals are merely end states or external standards that energize and direct behavior. Such mechanistic assumptions about the nature and connotations of goals do not provide insight into what contributes to differences in the goals that individuals adopt, whether assigned by the experimenter, set by the participant, and/or spontaneously generated in "no goal" or vague goal conditions. Moreover, such a perspective on the purpose of goals does little to inform us about the implications of such goals for variation in ensuing motivation-related responses and achievement behaviors.

Although Latham and Locke considered goal setting to be a "motivational technique" that is effective, a conceptual (or indeed, to some degree, even empirical) consideration of motivational processes is sorely missing in the goal-setting literature stemming from Locke's contributions. For example, commitment is assumed to be an important moderator of the goal setting–performance relationship in Locke's framework, but what contributes to variability in commitment to goals, both at the onset and throughout the goal-setting enterprise? Adequate ability is held to be fundamental for translating goals into desired performance, but goal-setting theory does not consider potential discrepancies in how individuals judge their level of competence or suggest when or why low ability would prove problematic. In Locke's goal-setting theory, difficult goals are presumed to foster effort, and this enhanced exerted effort is held to be a central way by which difficult goals contribute to superior performance. It is important to recognize, however, that people do not always try when the situation or task becomes challenging. Indeed, they may even intentionally "hold back" in such circumstances. Why might this occur? In short, such broader motivational issues, which have important implications for the degree to which discrete goals would lead to better performance and enhanced task or activity engagement, are not adequately addressed in Locke's initial conceptualization or later treatises on goal setting.

2. THE MEANING OF GOALS

A contemporary theory of motivation that emphasizes the meaning of personal goals and the consequences of this meaning for corresponding motivational processes

and achievement behaviors is achievement goal theory (AGT). Achievement goals are held to reflect different ways of construing one's competence in terms of the demands and standards placed on the individual and the different ways of subjectively defining successful performance and performance progress. It is assumed, based on AGT, that the demonstration of high competence (or avoidance of demonstrating low competence) is at the heart of striving for achievement and serves as the core of people's views regarding goal accomplishment. In the prevailing achievement goal conceptualizations that have dominated sport research over the past 10 to 15 years, two major achievement goals are deemed to be operating. One is a task-involved goal, which entails an emphasis on demonstrating self-referenced competence. In this case, witnessing personal improvement, learning, and/or task mastery provide a sense of subjective success. If centered on an ego-involved goal, however, revealing superior competence relative to others is fundamental to perceptions of goal accomplishment. Perceptions of adequate (or inadequate) ability are normatively referenced when ego-involved and, as a result, performance attempts stemming from high effort and reflecting personal progress are insufficient for feeling successful. Thus, although AGT considers the level of one's perceptions of ability (high/low) to have motivational significance, the hallmark of this framework is that we must bear in mind how the person judges whether he or she has exhibited high versus low ability when engaged in a task.

Drawing from AGT, it is proposed that a focus on task-involved goals, regardless of whether the person in question is confident or perceives her or his ability to be less than outstanding, will result in a positive and adaptive achievement pattern. That is, a task-involved person is expected to be more likely to perform up to her or his ability level, try hard and be invested in the activity, prefer challenging tasks, and persevere, even when the circumstances become difficult. As long as things are going well (i.e., the person is experiencing successful goal accomplishment and perceives his or her ability to be high), an individual striving to meet ego-involved goals should be committed to the activity at hand and perform well. In this situation, the person would not shy away from challenge and would be willing to work—as long as that level of exerted effort does not take the spotlight off or lead to questions regarding her or his level of superior competence. Short- and especially long-term motivational problems are hypothesized when there is a centering on ego-involved goals and the individual has reached the conclusion that her or his competence is inferior relative to relevant others. Such a person would be expected to reduce the commitment to the activity and rescind effort (perhaps in an initial attempt to "save face," but eventually, if the situation does not change, as a step toward withdrawal from this endeavor). Optimally challenging goals would be especially worrisome for an individual who is trying to meet ego-involved standards and who has doubts about her or his competence. Rather, the preference would be to work toward easy goals (so the probability of failing is reduced) or set goals that are much too difficult (so that failure can be attributed to task difficulty rather than one's inadequacy).

As Hall and Kerr eloquently argued, a consideration of achievement goals provides insight into what may be the meaning of assigned goals in typical goal-setting protocols and what lays the foundation for the personal goals that study participants might set in goal-setting investigations. Variations in achievement goals are also relevant to what type of feedback is sought and how the feedback provided is understood and used in studies of goal-setting effectiveness. Achievement goals are pertinent to the extent that competition is created among study participants in goal-setting studies, the reasons for competing, and how the competitive experience is construed (e.g., as a challenge or a threat). Finally, as achievement goals are assumed to act like an interpretive lens impacting how we view the task and ourselves in relation to the task (e.g., does our task performance have implications for self worth?), differences in achievement goal emphases should correspond to variations in commitment to the discrete goals that are set and the effort given to reach those standards, especially over time. The focus placed on task and ego goals also has repercussions for individuals' level of anxiety, strategy use, degree of confidence or self-efficacy, and other psychological and emotional mechanisms that lead to differential performance. In short, achievement goals color our approach to and the behavioral, cognitive, and affective consequences of the setting and striving for discrete performance standards.

2.1. Basic Tenets of the Achievement Goal Framework

Achievement goal theory holds that, when performing achievement-related tasks, individuals can fluctuate in their state of involvement directed toward task or ego

goals. That is, they can be more or less task- and ego-involved at any point during task engagement. The likelihood of being task- and/or ego-involved is assumed to be influenced by dispositional tendencies regarding these states of goal involvement; i.e., people (once they have reached a particular level of cognitive development in late childhood and beyond) can vary in their degree of task and ego orientation. It is important to point out that these orientations tend to be orthogonal. Thus, with specific reference to goal setting, we are interested in how individuals who are high task/high ego, high task/low ego, low task/high ego, and low task/low ego might differ in the goals set, their processing of goal-related feedback, and the corresponding motivational processes and resulting behaviors. In the sport domain, two instruments with established validity and reliability have been developed to assess dispositional goal orientations, namely the Task and Ego Orientation in Sport Questionnaire and the Perceptions of Success Questionnaire.

Whether someone is striving to meet task- and/or ego-involved goal criteria during the goal setting process is also assumed to be dependent on the situationally and task-emphasized goals or the perceived motivational climate. In particular, the tasks involved and social environments surrounding individuals as they undergo goal setting vary in the degree to which they are more or less task- and/or ego-involving. With respect to the goal-setting process, these overriding climates are created by the type and level of the goals assigned, the nature of the feedback given, the evaluation criteria used to judge goal accomplishment, and the manner in which that goal accomplishment is recognized. If goal setting was undertaken in a task-involving climate, the goals set would be primarily improvement-focused and effort-dependent (i.e., process and performance goals), the feedback provided would be informational, self-referenced, and task-centered, the person himself/herself could be involved in the evaluation of goal accomplishment, and successful goal completion would be primarily tied to intrinsic satisfaction. The individual's level of commitment, perspective on goal challenge, and input would be strongly recognized in a task-involving goal-setting program.

Goal setting in an ego-conducive manner, on the other hand, would entail a preoccupation with normatively referenced goals (the majority would probably be outcome-oriented), the feedback given would be more judgmental and marked by social comparison, evaluation of goal accomplishment would make the person particularly aware of and concerned about his or her relative ability, and the sources and nature of the recognition offered when goals are met (or not met) would help couple the individual's sense of self to successful (or unsuccessful) goal completion. Further, personal volition and the participant's contribution to the process would most likely take a back seat in an ego-involving goal-setting program.

Pertinent to sporting contexts, assessments have been developed to tap the perceived motivational climate created by the coach (e.g., the PMCSQ-2) and parents (e.g., the PIMCQ), in the view of athletes. Researchers have also pulled from the TARGET taxonomy formulated by Epstein to analyze the dimensions underlying the climates manifested in sport. The particular structures encapsulated in this taxonomy are the nature of the task, source of authority, bases and form of recognition, structure and focus of groups in the setting, criteria for evaluation, and use of time.

2.2. Achievement Goals, Discrete Goals, and Goal-Setting Styles

One source of confusion concerning goal setting in the achievement goal literature is the tendency to mistakenly equate task and ego goal orientations with differential goal-setting styles or discrete goal foci. As indicated above, contemporary work on goal setting acknowledges the existence of different types of discrete goals, i.e., process, performance, and outcome goals. Researchers and practitioners in sport psychology, when making their case that outcome goals and ego orientation should not be considered hazardous to achievement-related health, have implied that a pronounced ego orientation means that one predominantly aims to meet outcome goals. In contrast, it is suggested that a strong task orientation translates into a marked orientation toward process (and perhaps performance) goals. Other researchers and practitioners in the field have challenged these presuppositions. For example, Duda has argued that a person can interpret process-related, performance, and outcome standards in a task-involved as well as an ego-involved fashion.

According to Burton, the construct of goal-setting style refers to athletes' preferences regarding discrete goals during training or competition. Burton has proposed that there are three distinct goal-setting styles, namely performance-oriented, success-oriented, and failure-oriented. Performance-oriented athletes are assumed to have a pronounced task orientation.

Success- and failure-oriented athletes are held to be strongly ego-oriented. However, in the former instance, perceptions of ability are high and, in the latter, perceptions of ability are deemed inadequate. In Burton's competitive goal-setting model, athletes' dispositional goal orientations in concert with their perceived competence are hypothesized to result in the focus on particular discrete goals of varying difficulty (e.g., a performance orientation will likely lead to the setting of challenging performance goals, while a success orientation will most probably result in an emphasis on moderately difficult outcome goals). Once the discrete goal is selected, the direct and indirect effects embedded in Locke and Latham's framework should then come into fruition. Thus, achievement goals are presumed to primarily come to bear in terms of goal selection in this model.

In addition to these conceptual distinctions between goal orientations, discrete goals, and goal setting styles, Hall and Kerr pointed out that there is no compelling evidence supporting the errantly assumed linkages between these constructs (e.g., that highly task-oriented athletes do not set outcome goals). Moreover, discussed above, empirical studies of the goal-setting practices of athletes indicate that many athletes set process, performance, and outcome goals and focus on these divergent goals for different reasons. The achievement goal framework would offer the following to discussions regarding the variety in and relevance of discrete goals operating in sport: the prediction and understanding of an athlete's cognitive, affective, and behavioral responses over time are primarily a function of whether he or she is task- and/or ego-involved, not whether the athlete is currently striving to meet a process-, performance-, or outcome-based standard. That is, what such diverse standards mean to the athlete and the potential significance of meeting (or especially not meeting) the standard fluctuates depending on achievement goal emphases. It is these latter appraisals that color ensuing motivational processes.

With respect to research and applied work, therefore, it is important to keep in mind that athletes' goal orientations offer insight into their concerns when approaching, engaged in, or concluding involvement with achievement tasks. These goal orientations also provide awareness into different perspectives on the connotations of goal accomplishment (or the failure to meet goals). They can be conceptualized as a type of cognitive schema that influence the discrete goals that we may set in the physical domain as well as every other facet of goal-setting and goal-striving processes. On the other hand, as stated by Hall and Kerr (2001) goals as considered in the goal-setting literature are "end states that regulate human action by specifying an aim or an objective standard for a specific task" (p. 227).

Perhaps it is best to illustrate these qualitative differences with a metaphor. Discrete performance standards are the star we envision visiting and the impetus for us to jettison into space. They are the dream or aim that comes closest to the landing on the star. Goal orientations, however, capture the reason for going into the "last frontier" in the first place. They guide the features and resiliency of the spaceship we decide to take, the characteristics of the long ride to that distant star, the reactions when we reach our destination, and the ramifications of such an achievement for the entire space program. That is, goal orientations and discrete goals are not one and the same.

2.3. Potential Mechanisms Influencing the Goal Setting–Performance Relationship

What might be the mechanisms by which achievement goals impact the goal setting– performance relationship? The contemporary sport psychology literature contains a plethora of studies examining the interplay between achievement goals and various cognitive, affective, and behavioral responses in the sport domain. It is beyond the scope of this article to adequately review and do justice to this body of work. As they seem particularly applicable to the goal-setting process, though, it is important to highlight that variation in achievement goals (whether dispositional and/or situationally induced) have been found to correspond to (1) perceptions of ability, (2) differential beliefs about the causes of success, (3) anxiety responses and ways of coping with stress, (4) learning and performance strategy use, and (5) preferences regarding performance feedback.

In terms of individuals' perceived competence, both task and ego orientation (in cross-sectional investigations) tend to be positively associated. The correlations in these studies, however, tend to be low (typically, r = .2–.3), which tells us that at any particular point in time, we can have highly task- and/or ego-oriented individuals who feel their competence to be high, moderate, or far from stellar. In research on the interdependence

between environmental achievement goal emphases and perceptions of ability, task-involving situational cues have been found to correspond to higher perceptions of ability. Ego-involving situational cues, however, seem more likely to lead to a deterioration of a sense of adequate competence, especially if perceptions of ability were already depleted or if the circumstances (e.g., experience of failure) are leading to questions about one's ability level. These findings make sense when one considers that a task goal emphasis means that one construes personal competence in a self-referenced way and centers on effortful attempts at task mastery and/or improvement. Such a perspective should contribute to a sense of demonstrated competence on numerous occasions. In the case of an ego goal emphasis, witnessing one's superiority with respect to others translates into high perceived ability. One's confidence to do what needs to be done would be more likely to "take some hits" when competence and success are construed in that manner.

Locke and Latham's theory of goal setting holds that a person must feel that he or she has what it takes for the setting of a discrete goal (especially a difficult one) to result in heightened performance. Whether one's capability is deemed to be sufficient to meet the standard set, it seems, is dependent on the achievement goal(s) manifested.

With respect to views about what leads to achievement, task orientation tends to be linked to the belief that hard work and training is essential. It is reasonable to suspect that such an outlook on the determinants of success should translate into a resilient tendency to not give up during the "failures" and continuously exert effort during the goal-setting process. Ego orientation, on the other hand, has been found to relate to the belief that the possession of high ability is necessary for success to occur. One cannot help but wonder, then, if someone is pursuing discrete goals in an ego-involved manner, what would be the impact of this corresponding belief when goal accomplishment has not been forthcoming?

A plethora of studies in the field of sport psychology have explored the anxiety–performance relationship. When stress levels go beyond what is optimal for the task and person at hand, performance debilitation ensues. Although much more work needs to be done on this topic, previous research has indicated that task orientation and a perceived task-involving climate tend to be negatively or unrelated to pre-competitive anxiety, whereas a positive association between dispositional or situationally induced ego goals (especially in the case of individuals with low perceived ability) and anxiety responses has received support. In terms of coping responses, the literature typically points to a positive relationship between task goals and problem-focused strategies (which, in the short- and long-term, have been found to be effective). In contrast, ego orientation and a perceived ego-involving climate have been found to correspond to emotion- or avoidance-focused coping (which may be deemed beneficial at the moment, but does not correspond to long-term indicators of coping effectiveness).

Finally, preliminary work indicates that the achievement goals that are adopted have relevance for strategy use and preferences regarding performance feedback. In the studies to date, task goals have been coupled with more adaptive training and competition strategies such as trying to understand the coach's instructions and engaging in deep-processing when trying to learn a motor task. When athletes are high in ego orientation and/or perceive the climate as highly ego-involving, they have been found to be more likely to view avoiding practice as a viable strategy and to employ superficial learning strategies. Cury and his colleagues have conducted a series of interesting investigations regarding the type of feedback desired following task completion. Their research suggests that highly ego goal-focused individuals who doubt their ability tend to reject task-related or objective performance-related feedback. Clearly, it is difficult to see such people effectively enhancing their perceived or actual ability level with such feedback preferences. Strongly ego-oriented individuals who have high perceived ability seem more likely to seek out normative information (i.e., how good am I compared to others?) and reveal a low preference for objective performance-related feedback. Although seemingly not as debilitating as the feedback preferences of their high ego/lower perceived ability cohorts, such inclinations regarding preferred feedback are probably not optimal for meeting complex and difficult goals in the sport setting.

3. ADDITIONAL CONSIDERATIONS AND FUTURE DIRECTIONS

In the view of Hall and Kerr (2001), the attractiveness of the achievement goal framework is that it enables researchers and practitioners "... to understand considerably more about the direction of behavior and about why athletes make particular motivational choices than is possible from considering Locke and Latham's (1990)

goal-setting theory alone . . . [which] limits us to understanding changes in intensity and persistence once discrete goals are set" (pp. 230–231).

Systematic research that attempts to integrate the constructs and tenets of achievement goal theory with the hypothesized goal-setting process is needed. In essence, the aim of such work would be to better inform us about when and why goal setting would be effective or ineffective, and what might be the best ways to implement goal setting in the sport domain. With respect to the potential mechanisms involved, one research direction would be to test variations in the proposed direct and indirect effects of goals on performance as proffered in Locke's goal-setting theory as a function of the achievement goals manifested. In terms of differences in achievement goal emphases, it would be interesting to consider people's goal orientations and the motivational climate in which they are operating as well as the interaction between these two factors. It would also be conceptually intriguing and practically important to examine the relationship of achievement goals to the nature and motivational implications of the process, performance, and outcomes goals set by athletes in both training and competition. While pursuing such lines of inquiry, the impact of variability in athletes' perceived ability on the interdependence between achievement goals and the goal-setting process should be determined.

In the discussion of achievement goals in this article, the emphasis has been placed on what have recently been termed dichotomous goal models. Such models distinguish competence-based striving in terms of a task (or mastery) and ego (or performance) focus. This approach was adopted for several reasons. First, the two-goal approach to achievement motivation has laid the foundation to date for the large majority of achievement goal research in sport. Second, existing discussions in the literature concerning distinctions between the constructs of achievement goals and discrete goals have been grounded in dichotomous conceptualizations. Finally, space constraints precluded the possibility of fully describing and developing the recently proposed trichotomous goal models and "2 × 2" achievement goal framework in terms of the present discourse. These latter frameworks differentiate approach achievement goals [goals directed toward the attainment of favorable judgments of competence, whether that competence is self-referenced or task-based (i.e., approach task) or normatively referenced (i.e., approach ego)] from avoidance goal perspectives [goals centered on avoiding unfavorable judgments of ability, whether those perceptions of ability are self-referenced or task-based (i.e., avoidance task) or normatively referenced (i.e., avoidance ego)]. As these multiple goals have been found to differentially link to motivation-related responses such as intrinsic interest, anxiety, and performance, incorporating the approach–avoidance distinction in forthcoming studies of achievement goals and goal setting should prove informative.

Duda made a distinction between the quantity and quality of athletes' motivation. The former refers to an athlete being energized at that point in time and performing adequately. The latter relates to whether (1) the athlete wants to and does invest over time, (2) the activity contributes to the athlete's physical, psychological, emotional, and moral growth and well-being, (3) the athlete usually is able to perform up to her or his ability, and (4) the athlete enjoys the activity and is self-determined. At best, goal-setting theory alone can provide an appreciation of whether motivation is high or low from a quantity perspective. Integrating goal setting within an achievement goal framework, however, promises insight into both aspects. In so doing, we emphasize other achievement behaviors (e.g., perseverance) along with performance. Most critically, though, we consider other motivational outcomes of significance besides achievement (e.g., the athletes' welfare). Finally, as pointed out by Hall and Kerr, we become cognizant of the fact that (and the reasons why) goal setting can help promote quality sport engagement or be maladaptive (i.e., contributing to the quantity of motivation but compromising the quality and/or eventually leading to an erosion of quantitatively high motivation as well). All in all, based on the tenets of AGT and related research, it seems that more task-involving goal setting is what will reap the greatest motivational gain in sport or any other achievement domain.

See Also the Following Articles

Arousal in Sport ■ Intrinsic and Extrinsic Motivation in Sport ■ Psychological Skills Training in Sport ■ Self-Confidence in Athletes

Further Reading

Ames, C. (1992). Classrooms: Goal, structures, and student motivation. *Journal of Educational Psychology*, *84*, 261–271.

Burton, D. (1989). Winning isn't everything: Examining the impact of performance goals on collegiate swimmers' cognitions and performance. *The Sport Psychologist*, *3*, 105–132.

Burton, D., & Naylor, S. (2002). The Jekyll/Hyde nature of goals: Revisiting and updating goal-setting in sport. In T. Horn (Ed.), *Handbook of research on sport psychology*, 2nd ed. (pp. 459–500). Champaign, IL: Human Kinetics.

Burton, D. (1993). Goal setting in sport. In R. N. Singer, M. Murphey, & L. K. Tennant (Eds.), *Handbook of research on sport psychology* (pp. 467–491). New York: Macmillan.

Burton, D., Naylor, S., & Holliday, B. (2001). Goal setting in sport: Investing the goal effectiveness paradox. In R. Singer, H. A. Hausenblas, & C. M. Janelle (Eds.), *Handbook of sport psychology* (pp. 497–528). New York: John Wiley.

Burton, D., Weinberg, R., Yukelson, D., & Weigand, D. (1998). The goal effectiveness paradox in sport: Examining the goal practices of collegiate athletes. *The Sport Psychologist*, *12*, 404–418.

Chidester, T. R., & Grigsby, W. C. (1984). A meta-analysis of the goal-setting performance literature. In J. A. Pearce, & R. B. Robinson (Eds.), *Academy of management proceedings* (pp. 202–206). Ada, OH: Academy of Management.

Duda, J. L. (1989). The relationship between task and ego orientation and the perceived purpose of sport among male and female high school athletes. *Journal of Sport and Exercise Psychology*, *11*, 318–335.

Duda, J. L. (1992). Sport and exercise motivation: A goal perspective analysis. In G. Roberts (Ed.), *Motivation in sport and exercise* (pp. 57–91). Champaign, IL: Human Kinetics.

Duda, J. L. (1996). Maximizing motivation in sport and physical education among children and adolescents: The case for greater task involvement. *Quest*, *48*, 290–302.

Duda, J. L. (1997). Perpetuating myths: A response to Hardy's 1996 Coleman Griffith Address. *Journal of Applied Sport Psychology*, *9*, 307–313.

Duda, J. L. (2001). Goal perspectives in sport: Pushing the boundaries and clarifying some misunderstandings: In G. C. Roberts (Ed.), *Advances in motivation in sport and exercise* (pp. 129–182). Champaign, IL: Human Kinetics.

Duda, J. L. (2004). Verifying the values and sparking the spirit: Research and practice on motivation in sport psychology. In P. Markula (Ed.), *Sharing joy, sharing pain: Feminist narratives on sport studies* (in press). New York: SUNY Press.

Duda, J. L., & Hall, H. K. (2001). Achievement goal theory in sport: Recent extensions and future directions. In R. Singer, H. Hausenblas, & C. Janelle (Eds.), *Handbook of sport psychology,* 2nd ed. (pp. 417–443). New York: John Wiley & Sons.

Duda, J. L., & Nicholls, J. G. (1992). Dimensions of achievement motivation in schoolwork and sport. *Journal of Educational Psychology*, *84*, 290–299.

Duda, J. L., & Whitehead, J. (1998). Measurement of goal perspectives in the physical domain. In J. L. Duda (Ed.), *Advances in sport and exercise psychology measurement* (pp. 471–484). Morgantown, WV: Fitness Information Technology.

Dweck, C. S. (1999). *Self-theories and goals: Their role in motivation, personality, and development*. Philadelphia, PA: Taylor & Francis.

Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychology*, *34*, 169–189.

Elliot, A. J., & Church, M. A. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology*, *72*, 1–15.

Elliot, A. J., & McGregor, H. A. (2001). A 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology*, *80*, 501–519.

Epstein, J. (1989). Family structures and student motivation: A developmental perspective. In C. Ames, & R. Ames (Eds.), *Research on motivation in education* (Vol. 3) (pp. 259–295). New York: Academic Press.

Fry, M. D. (2001). The development of motivation in children. In G. C. Roberts (Ed.), *Advances in motivation in sport and exercise* (pp. 51–78). Champaign, IL: Human Kinetics.

Hall, H. K. (1990). *A social-cognitive approach to goal setting: The mediating effects of achievement goals and perceived ability*. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.

Hall, H. K., & Byrne, A. T. J. (1988). Goal setting in sport: Clarifying recent anomalies. *Journal of Sport and Exercise Psychology*, *10*, 184–198.

Hall, H. K., & Kerr, A. W. (2001). Achievement goal research in sport: Pushing the boundaries and clarifying some misunderstandings. In G. Roberts (Ed.), *Advances in motivation in sport and exercise* (pp. 183–234). Champaign, IL: Human Kinetics.

Hardy, L. (1997). Three myths about applied consultancy work. *Journal of Applied Sport Psychology*, *9*, 277–294.

Hardy, L., Jones, G., & Gould, D. (1996). *Understanding psychological preparation for sport: Theory and practice for elite performers*. Chichester, UK: Wiley.

Harwood, C., Cumming, J., & Fletcher, D. (2004). Motivational profiles and psychological skill use within elite youth sport. *Journal of Applied Sport Psychology,* in press.

Jones, G., & Hanton, S. (1996). Interpretation of competitive anxiety symptoms and goal attainment expectancies. *Journal of Sport and Exercise Psychology*, *18*, 144–157.

Kim, M-S., & Duda, J. L. (2004). The doping process: Cognitive appraisals of stress, coping strategies, and coping effectiveness. *The Sport Psychologist*, *17*.

Kingston, K. M., & Hardy, L. (1997). Effects of different types of goals on processes that support performance. *The Sport Psychologist*, *11*, 277–293.

Locke, E. A. (1968). Toward a theory of task motivation and incentives. *Organizational Behavior and Human Performance*, *3*, 157–189.

Locke, E. A. (1996). Motivation through conscious goal setting. *Applied and Preventative Psychology*, *5*, 117–124.

Locke, E. A., & Latham, G. P. (1985). The application of goal setting to sports. *Journal of Sport Psychology*, *7*, 205–222.

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting and task performance*. Englewood Cliffs, NJ: Prentice Hall.

Locke, E. A., Shaw, K. N., Saari, L. M., & Latham, G. P. (1981). Goal setting and task performance: 1969–1980. *Psychological Bulletin*, *90*, 125–152.

Mento, A. J., Steel, R. P., & Karren, R. J. (1987). A meta-analytic study of the effects of goal setting on task performance: 1966–1984. *Organizational Behaviour and Human Decision Processes*, *39*, 52–83.

Newton, M., & Duda, J. L. (1999). The interaction of the motivational climate, dispositional goal orientations, and perceived ability in predicting indices of motivation. *International Journal of Sport Psychology*, *30*, 63–82.

Newton, M., Duda, J. L., & Yin, Z. (2000). Examination of the psychometric properties of the Perceived Motivational Climate in Sport Questionnaire-2 in a simple of female athletes. *Journal of Sport Sciences*, *18*, 275–290.

Nicholls, J. G. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological Review*, *91*, 328–346.

Nicholls, J. G. (1989). *The competitive ethos and democratic education*. Cambridge, MA: Harvard University Press.

Ommundsen, Y., & Roberts, G. C. (1999). Effect of motivational climate profiles on motivational indices in team sport. *Scandinavian Journal of Medicine and Science in Sports*, *9*, 389–397.

Roberts, G. C. (2001). Understanding the dynamics of motivation in physical activity: The influence of achievement goals on motivational processes. In G. C. Roberts (Ed.), *Advances in motivation in sport and exercise* (pp. 1–50.). Champaign, IL: Human Kinetics.

Roberts, G. C., Treasure, D. C., & Balague, G. (1998). Achievement goals in sport: The development and validation of the Perception of Success Questionnaire. *Journal of Sport Sciences*, *16*, 631–645.

Swain, A., & Jones, G. (1995). Effects of goal-setting interventions on selected basketball skills: A single-subject design. *Research Quarterly for Exercise and Sport*, *66*, 51–63.

Thill, E., & Brunel, P. (1995). Ego involvement and task involvement: Related conceptions of ability, effort, and learning strategies among soccer players. *International Journal of Sport Psychology*, *30*, 101–122.

Treasure, D. C. (1997). Perceptions of the motivational climate and elementary school children's cognitive and affective responses. *Journal of Sport and Exercise Psychology*, *19*, 278–290.

Treasure, D. C. (2001). Enhancing young people's motivation in youth sport: An achievement goal approach. In G. C. Roberts (Ed.), *Advances in motivation in sport and exercise* (pp. 79–100). Champaign, IL: Human Kinetics.

Tubbs, M. E. (1986). Goal setting: A meta-analytic examination of the empirical evidence. *Journal of Applied Psychology*, *71*, 474–483.

Weinberg, R. S., Bruya, L. D., & Jackson, A. (1985). The effects of goal proximity and goal specificity on endurance performance. *Journal of Sport Psychology*, *7*, 296–305.

Weinberg, R. S., Bruya, L. D., & Jackson, A. (1990). Goal setting and competition: A reaction to Hall and Byrne. *Journal of Sport and Exercise Psychology*, *12*, 92–97.

Weinberg, R. S., Burke, K. L., & Jackson, A. (1997). Coaches' and players' perceptions of goal setting in junior tennis: An exploratory investigation. *The Sport Psychologist*, *11*, 426–439.

Weinberg, R. S., Burton, D., Yukelson, D., & Weigand, D. (1993). Goal setting in competitive sport: An exploratory investigation of practices of collegiate athletes. *The Sport Psychologist*, *7*, 275–289.

Weinberg, R. S., Burton, D., Yukelson, D., & Weigand, D. (2000). Perceived goal setting practices of Olympic athletes: An exploratory investigation. *The Sport Psychologist*, *14*, 279–295.

Weinberg, R. S., Butt, J., & Knight, B. (2001). High school coaches' perceptions of the process of goal setting. *The Sport Psychologist*, *15*, 20–47.

Weinberg, R. S., Butt, J., Knight, B., & Perritt, N. (2001). Collegiate coaches' perceptions of their goal setting practice: A qualitative investigation. *Journal of Applied Sport Psychology*, *13*, 374–398.

Weinberg, R. S., & Weigand, D. (1993). Goal setting in sport and exercise: A reaction to Locke. *Journal of Sport and Exercise Psychology*, *15*, 88–96.

Weinberg, R. S., & Weigand, D. (1996). Let the discussions continue: A reaction to Locke's comments on Weinberg and Weigand. *Journal of Sport and Exercise Psychology*, *18*, 89–93.

White, S. A. (1996). Goal orientations and perceptions of the motivational climate initiated by parents. *Pediatric Exercise Science*, *8*, 122–129.

Wood, R. E., Mento, A. J., & Locke, E. A. (1987). Task complexity as a moderator of goal effects: A meta-analysis. *Journal of Applied Psychology*, *72*, 416–425.

Woodman, T., & Hardy, L. (2001). Stress and anxiety. In R. Singer, H. Hausenblas, & C. Janelle (Eds.), *Handbook of sport psychology*, 2nd ed. (pp. 290–318). New York: John Wiley & Sons.

Group Counseling

Janice L. DeLucia-Waack
University at Buffalo, State University of New York, Buffalo, New York, USA

1. Introduction
2. Definition of Group Counseling
3. Therapeutic Factors Theory and Research
4. Effective Interventions Theory and Research
5. Considerations
6. Conclusion
Further Reading

This article provides a theoretical and empirical rationale for successful group counseling and therapy along with precise definitions. The literature is reviewed with regard to therapeutic factors and effective interventions in group counseling. Considerations for facilitating successful counseling groups are suggested.

GLOSSARY

group counseling A small group intervention that uses therapeutic factors and specific interventions to address personal and interpersonal problems and to promote personal and interpersonal growth and development (e.g., a women's group).

group therapy A small group intervention that uses therapeutic factors and specific interventions to address personal and interpersonal problems and to remediate perceptual and cognitive distortions or repetitive patterns of dysfunctional behavior (e.g., an eating disorders group).

psychoeducational and guidance groups A small group intervention that uses group dynamics and principles to teach skills in a preventive way (e.g., a friendship skills group).

task and work groups A small group activity that uses organizational and community dynamics to facilitate the goals of a group (e.g., a book club).

therapeutic factors An element of group therapy that contributes to improvement in a patient's condition and is a function of the actions of the group therapist, the other group members, and the client himself or herself.

1. INTRODUCTION

Current group counseling theory and practice has its roots in two divergent events: Sigmund Freud's famous Wednesday night meetings at his house in Vienna and John Pratt's treatment of tuberculosis patients in group "classes" in the United States. Many theoretical approaches and applications have emerged from these early beginnings to create an efficient and empirically validated treatment in psychology today. As research methods became more precise and more interest in studying the effectiveness of groups accrued, the focus moved from showing that groups were more effective than no treatment at all to establishing that groups for certain populations and problems were as effective as, or even more effective than, other active treatments. As research, theory, and practice developed in the field, the pertinent question changed from "Are groups effective?" to "What kinds of groups are effective for what types of people with what problems?"

©2004 Elsevier Inc.
All rights reserved.

2. DEFINITION OF GROUP COUNSELING

Toward the goal of answering the question of what kinds of groups are most effective for certain types of people with certain types of problems, the Association for Specialists in Group Work (ASGW) in 2000 delineated four types of groups based on their goals and interactional processes—task/work, psychoeducational/guidance, counseling, and therapy groups—to aid in the selection of the appropriate types of groups for different populations (e.g., age groups) with different goals (e.g., combating depression, learning social skills, preventing eating disorders). According to ASGW, counseling groups "address personal and interpersonal problems of living and promote personal and interpersonal growth and development," whereas therapy groups "address personal and interpersonal problems of living, remediate perceptual and cognitive distortions or repetitive patterns of dysfunctional behavior, and promote personal and interpersonal growth and development." Examples of counseling and therapy groups include general interpersonal groups, training groups for students learning to be counselors or therapists, and groups directed at the amelioration of specific problems such as depression, eating disorders, and sexual abuse. Both counseling and therapy groups, as defined by ASGW, use "group-based cognitive, affective, behavioral, or systemic intervention strategies." Although by definition counseling and therapy groups are different in practice, these terms are often used interchangeably and are often difficult to differentiate in practice.

In contrast, task/work groups focus on "efficient and effective accomplishment of group tasks" and psychoeducational/guidance groups focus on skill development to prevent problems, according to ASGW. Both of these groups use "group-based educational and developmental strategies." Examples of task/work groups include committees, clubs, and classrooms that have common goals to work toward, whereas psychoeducational/guidance groups teach specific skills and coping strategies such as anger management, social skills, self-esteem, assertiveness, and making friends.

In further contrast, counseling and therapy groups tend to be of longer duration and smaller in size than do psychoeducational groups, the latter of which are often only 6 to 10 sessions and may consist of 10 or more members. Whereas task/work and psychoeducational/guidance groups tend to use educational strategies and practice to teach new skills, the emphasis in counseling and therapy groups is on the use of group process and dynamics to help members identify interpersonal difficulties, identify new ways of thinking and behaving to solve interpersonal problems, practice new behaviors, get feedback from other group members, and learn coping skills from other group members. The major premise inherent in counseling and therapy groups is that group members eventually act in the group in ways similar to how they behave in their relationships with others (e.g., significant others, friends, co-workers). Thus, as behaviors are enacted, group members can evaluate (with the help of other group members) the effectiveness of these behaviors and their impact (both positive and negative) on group members' relationships. In addition, group members then have the opportunity to practice and try out new behaviors and to receive feedback from others about the perceived impact of the new behaviors.

3. THERAPEUTIC FACTORS THEORY AND RESEARCH

Over the years, many group researchers and practitioners have attempted to explain the process of group counseling and therapy by describing 11 therapeutic factors that operate in groups. Although Yalom was not the first researcher or theorist to use the concept of therapeutic factors, his work has had an enormous influence on group theory and research. In 1975, Yalom's compendium of 11 therapeutic factors was built on earlier writings of group theorists and practitioners and on his research on important change mechanisms in group therapy. In 1985, Bloch and Crouch defined a therapeutic factor as "an element of group therapy that contributes to improvement in a patient's condition and is a function of the actions of the group therapist, the other group members, and the client himself." Table I provides the definitions of these 11 therapeutic factors.

Not only did Yalom define and describe 11 therapeutic factors, he also developed a research paradigm that has had a major influence on how researchers examine therapeutic factors in group settings. In a now classic study, Yalom and colleagues in 1973 developed a 60-item therapeutic factors Q-sort by writing five items that corresponded to each of the 11 factors described by Yalom (12 factors were used because interpersonal learning was split into interpersonal learning [input] and interpersonal learning [output] categories), asking group members to put items in

TABLE I
Descriptions of Therapeutic Factors

Therapeutic factor	*Definition*
Instillation of hope	Member recognizes other members' improvement and that group can be helpful; member develops optimism for his or her own improvement
Universality	Member perceives that other members share similar feelings or problems
Imparting information	Advice giving by therapist or fellow members
Altruism	Member gains a positive view of himself or herself through extending help to others in group
Corrective recapitulation of primary family group	Member experiences the opportunity to reenact some critical familial incident with members of the group in a corrective manner
Development of socializing techniques	Group provides member with an environment that allows member to interact in a more adaptive manner
Interpersonal learning (input)	Member gains personal insight through other members' sharing of their perceptions of the member
Interpersonal learning (output)	Group provides member with an environment that allows member to interact in a more adaptive manner
Cohesiveness	Feeling of togetherness is provided and experienced by the group
Catharsis	Member releases feelings about past or here-and-now experiences; this release leads to member feeling better
Existential factors	Member ultimately accepts that he or she must take responsibility for his or her own life
Imitative behavior	Member learns through the observation of others' learning experiences

categories ranging from "least helpful to me in group" to "most helpful to me in group." From these data, a ranking of the relative importance of the therapeutic factors, as perceived by successfully treated clients, was made from most important to least important as follows:

1. Interpersonal learning (input)
2. Catharsis
3. Cohesiveness
4. Self-understanding
5. Interpersonal learning (output)
6. Existential factors
7. Universality
8. Instillation of hope
9. Altruism
10. Family reenactment
11. Guidance
12. Identification

Yalom's paradigm for examining the importance of the therapeutic factors has spawned a large number of studies of therapeutic factors in groups. Several measures have been developed to assess therapeutic factors in groups (e.g., Critical Incidents Questionnaire, Therapeutic Factors Scale, Curative Factors Scale, Therapeutic Factors Inventory). Group researchers have now amassed a number of studies that identify clients' ranking of the therapeutic factors across a number of different settings and patient populations. Studies over the years have assessed ratings of therapeutic factors and yielded inconsistent results. These inconsistencies may be due to differences in the groups surveyed or in the methods used to assess therapeutic factors. In an effort to make sense of the inconsistent findings, Kivlighan and Holmes used a cluster analysis of studies that rated the importance of therapeutic factors. Four clusters emerged from the analysis based on the therapeutic factors that were highest and lowest in each cluster. Cluster 1 (affective insight groups) rated acceptance, catharsis, interpersonal learning, and self-understanding as the most valued factors. These groups also rated both guidance and vicarious learning as relatively unimportant. Cluster 2 (affective support groups) rated acceptance, installation of hope, and universality as the most important group factors. Similar to Cluster 1, members of the groups in Cluster 2 rated guidance and vicarious learning as relatively unimportant. Cluster 3 (cognitive support groups) rated vicarious learning and guidance as highly valued while rating self-understanding much

lower. Cluster 4 (cognitive insight groups) rated interpersonal learning, self-understanding, and vicarious learning as the most valued therapeutic factors. The topic of therapeutic factors continues to be intriguing to researchers and practitioners alike, especially as related to the questions of importance in differential stages of group development and for different types of groups.

4. EFFECTIVE INTERVENTIONS THEORY AND RESEARCH

When leaders employ a given intervention during the course of a group session, there is the implicit assumption that their behavior will have a positive impact on individual members and/or on the processes and outcomes of the group. Recent research findings have indicated that group dynamics and outcomes are influenced by factors such as group leaders' interaction styles, personal characteristics, and attitudes. Researchers have found that leaders who are less controlling, exhibit warmth/caring, and set and reinforce clear norms are more likely to have cohesive groups. Increased cohesiveness, in turn, has been found to be positively related to a variety of group treatment outcomes. Morran and colleagues highlighted 10 specific group interventions with research and clinical support.

The first category of interventions, protecting group members and promoting safety, includes protecting, blocking, and supporting. These interventions are typically most useful in the leader's efforts to provide a group climate that is conducive to trust, openness, and cohesion. Research has indicated that members who are dissatisfied with their group experience often implicate the group leader for negligence in providing adequate protection. Sometimes, protecting can be indirect and might include member selection/exclusion procedures, the establishment of appropriate group norms, or the modeling of caring for group members. More direct protecting interventions might take the form of intervening to stop a member who is self-disclosing too much or at a level that is significantly more intimate than the that of the rest of the members. Blocking is a specific type of protection that is used to stop a member from storytelling, rambling, or inappropriately probing, gossiping, or invading the privacy of others. Research suggests that group members frequently attribute their damaging experiences to undue confrontation, criticism, and pressure to self-disclose by other members. Thus, it is essential that the group leader use blocking to protect potentially vulnerable group members. Supportive interventions help to reassure, reinforce, and encourage members to participate in the group and try out new behaviors. One caution is that premature or excessive supporting interventions may foster dependency on the leader and be detrimental to the individual member and the group. However, when timed appropriately, supporting interventions can provide the encouragement needed for a group member to become involved in the group interaction.

The second category of interventions, energizing and involving group members, includes drawing out, modeling, linking, processing, interpreting, self-disclosing, and providing feedback. These interventions are most often used to stimulate forward progress, increase member participation, and enhance interpersonal learning. Drawing out occurs when the leader directly invites comments or involvement to encourage participation from members who find it hard to share with others or with those who share on a surface level but avoid deeper issues. The leader should encourage participation while still allowing members to choose their own levels of sharing. Modeling occurs when the leader demonstrates the skills, attitudes, and other characteristics he or she hopes to engender within group members, including respect and caring for others, appropriate self-disclosure, feedback (both giving and receiving), and openness. Many group theorists suggest that modeling is a common factor of nearly all interventions because members naturally observe group leaders and other members and tend to imitate what they see demonstrated. Research has found that behaviors displayed by the therapist, including interpersonal behaviors, feedback delivery and acceptance, and "here-and-now" communications; lead to increases in those same behaviors by group members. As a whole, studies suggest that leaders can use modeling interventions to shape positive group dynamics; however, care should be exercised to avoid the unintentional modeling of undesirable behaviors. Linking connects what one group member is saying or doing with the concerns of one or more other members. Linking has been shown to be particularly useful in promoting member interaction, cohesion, trust, and universality among group members.

Processing as an intervention occurs when the group leader and members capitalize on significant happenings to help members reflect on the meanings of their experiences; better understand their own thoughts,

feelings, and actions; and generalize what is learned to situations outside of the group. Simply experiencing events in the group is not sufficient for growth and must be augmented by processing that provides a framework for retaining, integrating, and generalizing the experience. Research findings have identified a positive relationship between higher levels of group processing and greater outcome gains related to both group productivity and individual member achievement and suggest that processing should go beyond emotional insight and also focus on cognitive dynamics or meaning attributions to help members frame and understand their emotional experiences. Effective processing of critical incidents requires that members engage in a process of sharing and exploring experiences among themselves to make sense of what happened, establish what they have learned from the incidents, and begin to plan and practice new behaviors to be used outside of group sessions.

Interpreting involves offering possible explanations for events, behaviors, thoughts, and feelings at a deeper level to be considered as part of the change process. Group members typically experience interactions that provide the opportunity to learn more about themselves. However, such insight may be difficult for members to acquire without some cognitive framework being provided by the leader. In such situations, the leader can assist group members by providing tentative interpretations that can be discussed and considered in the group. Research suggests that leader interpretations help members to integrate complex personal and group-related events, thereby encouraging their investment in the group experience and facilitating generalization from group experiences to personal experiences outside of the group. Interpretations about members' impact on the environment and their patterns of behavior are most associated with member change.

Leader self-disclosure is an intervention whereby the therapist reveals his or her personal feelings, experiences, or here-and-now reactions to group members. As a general rule, the leader should self-disclose only when the information or reaction is directly related to what is happening in the group, models a behavior for members, and/or is expected to result in a direct benefit to the group.

5. CONSIDERATIONS

Much research has been conducted about counseling groups, but there is still much more that needs to be known about how to facilitate groups effectively. Several considerations are important in leading effective counseling groups. First, sufficient training is needed in the type of group to be facilitated. ASGW provides specific guidelines for the training and supervision of task, psychoeducational, counseling, and therapy groups. Research suggests that when group therapists lack supervision, they are not able to identify mistakes and generate new plans of action; instead, they become stuck in a cycle of repeated ineffective interventions. Co-leadership is also recommended as a training tool and generally to enhance group effectiveness. It is suggested that co-therapy is better for a variety of reasons such as efficiency (e.g., the group can still continue if one leader is sick or unavailable), effectiveness (e.g., two leaders can track group members more easily), supervision and training of new group therapists and as a way in which to model healthy communication patterns (e.g., male–female interactions) between co-leaders in front of the group members. It is also theorized that co-leaders who model appropriate interaction between themselves, such as resolving disagreements and compromising, will have group members who demonstrate better skills in these areas.

Second, diversity and multicultural issues groups are inherent in groups and are essential to consider. Group leader competence is critical, specifically multicultural group counseling skills. In addition, knowledge of other cultural worldviews and values, and of how these may inhibit or increase a potential group member's willingness to participate and behave in the group, is essential. Acculturation is noted as an important variable in predicting the behavior of potential group members. Effective groups must help members to understand themselves and others as individuals within the context of their culture and must choose interventions and methods of change based on the interplay between the individuals and their worldviews.

Third, there is support for conducting groups that are homogenous in age and gender due to the impact of cognitive development and different life tasks across age groups on the purpose and content of the groups. The structure of a group, types of exercises, focus on process versus content, and group goals all are determined by the abilities and needs of the group members.

Fourth, preventive and brief interventions can be effective and may ameliorate problems before they become more serious. Thus, psychoeducational groups may be useful, as may counseling and therapy groups.

6. CONCLUSION

Group counseling has been shown to be effective with various populations, in different settings, and for specific problem areas. Research clearly shows that group goals must be clearly defined, the leader must have specific leadership training and skills, and the leader must take into consideration group members' expectations about the group, willingness to participate, and cultural expectations and values when designing a group and implementing specific interventions.

See Also the Following Articles

Groupthink ■ Intergroup Relations and Culture ■ Interpersonal Behavior and Culture

Further Reading

Association for Specialists in Group Work. (2000). Association for Specialists in Group Work: Professional standards for the training of group workers. *Journal for Specialists in Group Work*, *25*, 327–354.

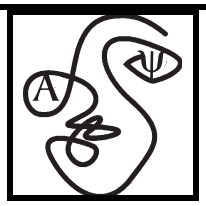
Barlow, S. H., Fuhriman, A. J., & Burlingame, G. (in press). The history of group counseling and psychotherapy. In J. L. DeLucia-Waack, D. A. Gerrity, C. R. Kalodner, & M. T. Riva (Eds.), *Handbook of group counseling and psychotherapy*. Thousand Oaks, CA: Sage.

Bloch, S., & Crouch, E. (1985). *Therapeutic factors in psychotherapy*. Oxford, UK: Oxford University Press.

Kivlighan, D. E., Jr., & Holmes, S. E. (in press). The importance of therapeutic factors in counseling groups: A typology of therapeutic factors studies. In J. L. DeLucia-Waack, D. Gerrity, C. R. Kalodner, & M. Riva (Eds.), *Handbook of counseling and group psychotherapy*. Thousand Oaks, CA: Sage.

Morran, D. K., Stockton, R., & Whittingham, M. (in press). Effective leader interventions for counseling and therapy groups. In J. L. DeLucia-Waack, D. Gerrity, C. R. Kalodner, & M. Riva (Eds.), *Handbook of counseling and group psychotherapy*. Thousand Oaks, CA: Sage.

Yalom, I. D. (1995). *The theory and practice of group psychotherapy* (4th ed.). New York: Basic Books.



Group Dynamics in Sport

Packianathan Chelladurai
Ohio State University, Columbus, Ohio, USA

1. Introduction
2. Definition of a Sport Team
3. Task Types in Sport
4. Measures of Group Performance
5. Cohesion in Sport Teams
6. The Multidimensional Model of Leadership
7. The Mediational Model of Leadership
8. Motivational Climate
Further Reading

GLOSSARY

athletics Pursuit of excellence characterized by dedication and intensity.
cohesion Sticking together to achieve group tasks and satisfy social needs.
dependence The extent of interactions among members in the performance of the team task.
group dynamics The processes within a group that contribute to its success or failure as well as the scientific study of such processes.
leadership The interpersonal process of influencing members toward the achievement of team goals.
player reaction The attitudes toward the coach, the team, and the sport cultivated as a function of leader behavior.
sport Pursuit of pleasure characterized by moderation and generosity.
sports team Two or more individuals sharing a common group identity, facing a common fate, and interacting with each other in interdependent activities in the pursuit of common goals.
variability The extent to which the environment changes during the performance of a task.

Group dynamics is a term used to refer to the processes within a group that contribute to its success or failure as well as the scientific study of such processes. In the context of sport, greater emphases have been placed on the processes of cohesion and leadership as they relate to performance and player affective reactions. This article outlines the multidimensional view of both cohesion and leadership and the measurement thereof. In addition, it summarizes research findings on the antecedents and consequences of cohesion and leadership.

1. INTRODUCTION

The popularity of sport around the world is evidenced in the economic worth of the sport industry. In the United States alone, the sport industry is estimated to be more than $160 billion, making it one of the 10 or 12 largest industries in the country. The sport industry is based on the degree to which people are identified with and committed to a sport, a team, and/or an athlete. Such commitment very often translates into heated arguments about the performances of one's team. Most discussions of the successes and failures of a sport team focus on how well one or more athletes performed and on the strategic decisions made by the coach. The debate would also extend to how well members of the

©2004 Elsevier Inc.
All rights reserved.

team get along with each other and the rapport between the coach and his or her players. These speculations of the "armchair" experts have been the substance of analysis and critique in group dynamics, a significant scientific endeavor within the broad realm of sport psychology. In this description of the field, the terms "sport groups" and "sport teams" are used interchangeably. It must be noted that group dynamics has been studied in the context of exercise groups as well. However, the treatment of the chosen topics of cohesion and leadership is confined to sport teams as groups.

2. DEFINITION OF A SPORT TEAM

In 1998, Carron and Hausenblas defined a sport team as follows:

> a collection of two or more individuals who possess a common identity, have common goals and objectives, share a common fate, exhibit structured patterns of interaction and modes of communication, hold common perceptions about group structure, are personally and instrumentally interdependent, reciprocate interpersonal attraction, and consider themselves to be a group.

This definition incorporates the significant attributes of a group highlighted previously by other scholars. One attribute is that the members of the group have a common identity. Although having their unique names, uniforms, coaches, and support groups indicates a separate identity for every team, it is equally important that members accept that identity, share it, and consider themselves to be part of the group. The goals and objectives of sport teams may vary across levels of competition (e.g., winning in pursuit of excellence, learning and having fun in pursuit of pleasure). Another significant attribute is that the group members share a common fate. When a team loses, everyone on the team loses. When a team is disqualified for rule violations committed by some members, all members suffer. Members of a sport team, in sharing a common goal and fate, also have a common understanding of the way in which the team is structured, the manner and mode of communications within the team (including the technical terminology associated with the sport, e.g., what do "screening" in basketball and "stealing" in baseball mean?), and the terms used in the special plays unique to the team. Members also recognize that they are interdependent on each other to achieve the team goals. Given these attributes, members tend to develop a liking for each other (i.e., the interpersonal attraction).

2.1. Unique Attributes of Sport Teams as Groups

Sport teams differ from other conventional groups in several respects. One unique attribute of a sport team is the constant roster size as specified by the rules of the game and/or the league (e.g., 12 members on a volleyball team, with 6 players on the court at any one time). It is also possible that a league, or a group of teams in a sport, may jointly decide to reduce the maximum roster size for financial reasons or other exigencies.

Another unique attribute is that most activities of the groups are stipulated and controlled by the rules of the game. For example, only certain positions are permitted to advance the ball in football, use of hands is forbidden in soccer except by the goalkeeper, and holding the ball is not allowed in volleyball. Thus, teams in a given sport are similar in their inputs, their throughput processes are based on shared technology, and their expected outputs are identical.

The league to which a sport team belongs can also regulate the processes to a considerable extent. For example, the National Collegiate Athletic Association (NCAA) stipulates the maximum number of athletic scholarships permissible for an intercollegiate athletic team. It also regulates when a team can start practicing and how many games it can play. In addition, it specifies the scholastic requirements for athletes to be a part of the team. The list goes on, but the point is that sport teams are distinct from other forms of groups due to these extensive external controls over internal processes.

Another important attribute that sets sport teams apart from traditional groups is the presence of clear standards of effort and performance. Members of a sport team recognize that all of them should report to the practice sessions, follow the instructions of the coaches, and work strenuously during every session. It is also clearly understood that during competitions, all players should put forth their best efforts and coordinate their activities in pursuit of victory. In interacting with each other on both task and social spheres, members of a sport team establish and commit themselves to these norms of effort and productivity. In general, these features contribute to the sense of "groupness" among members of a sport team.

Another distinguishing feature of sport teams is the simultaneous presence of competition among the members of the team (particularly in team sports) and the need for cooperation to achieve team success. For instance, because only 5 members of a basketball team can be on the floor at any one time, the team's 12 members compete

with each other for starting status and/or playing time. At the same time, they cooperate with each other in coordinating their activities for performance success. Successful coaches are those who encourage competition among their players so that each one will try harder to be better than the other players and, at the same time, foster cooperation and coordination among them.

2.2. Purposes of Participation in Sport

One criterion for classifying sport teams is the purpose of participation in sport. In 1964, Keating distinguished between "athletics" and "sport." Sport is derived from the French word *desporter,* meaning a diversionary activity seeking pleasure for both participants. Athletics, derived from the Greek words *athlos* (i.e., a contest), *athlon* (i.e., a prize), and *athlein* (i.e., to contend for a prize), is essentially a competitive activity. In sport, the pleasure and fun can be maximized with a spirit of moderation and generosity. In contrast, athletics is characterized by a spirit of dedication, sacrifice, and intensity. Because the term "athletics" refers to one form of sport, namely track and field in the international context, Keating's other labels—"pursuit of pleasure" and "pursuit of excellence"—are used to denote these two distinct domains. This distinction parallels the one between recreational sport and competitive sport. Within the preceding broader classification, distinctions among youth sports, adult sports, scholastic and collegiate sports, community sports, and professional sports are also commonly made.

3. TASK TYPES IN SPORT

Beyond the distinction between pursuit of excellence and pursuit of pleasure, sport teams do differ in the kinds of tasks in which they are engaged. Two significant attributes have been used to distinguish among sport tasks: dependence and variability.

3.1. Dependence

Based on the criterion of dependence, sport tasks are categorized as independent, coactively dependent, proactively–reactively dependent, and interactively dependent. An independent task is one in which an individual can complete the task without interacting with other members of the team (e.g., high jumping, discuss throwing, individual events in swimming, badminton singles). Individuals in coactively dependent tasks depend on a common but external source for initiation and/or control of their activities. In general, members in a coactively dependent team perform more or less similar tasks, and their collective contributions determine the group's success (e.g., rowing, synchronized swimming). Reactively–proactively dependent tasks are those in which one team member initiates a task and depends on another team member to complete the task. Conversely, the second person depends on the first person to initiate the task (e.g., football quarterback and receiver, successive runners in a four-person track relay team). Finally, interactively dependent tasks require the team members performing various assignments to interact with each other effectively to complete the group task (e.g., basketball, soccer, hockey).

Sports such as soccer, volleyball, and football are characterized by a high degree of interdependence (i.e., members need to interact with each other during the competitions), whereas sports such as track and field, swimming, and diving are low on this attribute (i.e., performers do not interact with each other in performing their respective tasks). In common parlance, sports from the former set are called "team" sports and those from the latter set are called "individual" sports. Yet coaches, athletes, the public, and the media do conceive of "teams" in individual sports (e.g., the national swimming team, a collegiate track and field team). In general, several athletes in individual sports such as swimming are usually placed under one coach, who tends to organize practice sessions for all athletes at the same specified time and place. During the practice sessions, athletes interact with each other both on a task basis (e.g., competing with each other, setting the pace for each other) and on an interpersonal basis (e.g., encouraging each other, discussing personal matters). From this perspective, athletes on a swimming team are members of a group, and the elements discussed here would be equally applicable to athletes in team sports and those in individual sports.

3.2. Variability

Another characteristic distinguishing sports tasks is variability (i.e., the degree of environmental changes requiring participants to adjust to these changes). In high jumping (i.e., a low-variability task), an individual initiates the performance and executes the jump in a relatively stable environment (i.e., the uprights and crossbar are fixed). In basketball rebounding, the

actions are triggered by constantly varying environmental stimuli (e.g., the timing of a shot, the location of the rebound, the location of teammates and opponents).

4. MEASURES OF GROUP PERFORMANCE

In sport, the performance measures are very often contaminated by factors such as random chance (e.g., the bounce of the ball), an opponent's outstanding performance, strategic choices made by the team or coach, and bad calls made by officials. This section outlines the issues associated with performance measures in sport.

4.1. Subjective Evaluation

One issue in the measurement of sport performance relates to the subjective evaluation of performance in some sports such as diving, figure skating, and gymnastics. Because of the fallibility of subjective judgments, it is customary to take the averages of the ratings of a specified number of judges to give a score to a performance. Even in sports such as soccer, basketball, hockey, and football, the calls or rulings of referees and officials may be called into question. For example, when two players collide in basketball, a referee may call it a charge (i.e., an offensive foul), whereas a coach may view it as a block (i.e., a defensive foul). The subjective nature of these decisions can provoke heated arguments and catcalls during competitions. The new rules in professional football permit a coach to question a specific call made by an official (e.g., whether a receiver was in-bounds when he caught the ball). The chief referee is bound to look at the official replays and decide whether the official made the correct call or not. It must be noted that the subjective element of a decision may at times be masked by the so-called objective measures. For instance, in football, officials take extraordinary efforts to "bring in the chain" and carefully verify whether the ball crossed the requisite 10 yards for a first down. All of these efforts at objective measurement are based on the subjective judgment of a single referee on where the ball hit the ground when the play ended.

4.2. Zero-Sum Nature of Competitions

Another issue relates to the zero-sum nature of sports competitions. That is, for every winner, there is a loser. To alleviate the problem, the win–loss record over an entire season is used as a measure of performance. Another issue is that a basketball team that loses by a score of 100 to 30 is given the same standing as a team that loses by a score of 100 to 99. One way in which to address this issue is to use the point differential (i.e., the difference between points scored for and against the team in a contest) or the ratio of the final score (of the two teams in a contest) as a measure of performance.

4.3. Team Performances in Independent Sports

Another issue is with the measurement of group performance where members are involved in individual tasks such as archery, badminton, boxing, tennis, golf, swimming, and track and field. For instance, in the international Davis Cup tennis competitions, there are four singles matches and one doubles match. The national team that wins the majority of these five matches is declared the winner. This format of deciding on team performance in individual sports is quite common across all levels of competition.

4.4. Perceptual Measures of Performance

Given that performance measures may be contaminated by several factors, some scholars have used the perceptions of athlete themselves regarding their personal performance and/or their teams' performance. The argument is that every athlete has set expectations for individual- and team-level performances and is aware of the efforts that he or she, and the team as a whole, had put forth toward achieving those expectations. Therefore, the athletes are the best judges as to whether they had attained their goals. By the same token, they could also be asked about the improvements they had made. These perceptual measures are ideal if the focus of the endeavor is on athletes' development and their affective reactions.

5. COHESION IN SPORT TEAMS

While the study of group dynamics in sport has gained steam during the past two decades or so, the most notable and sustained efforts have focused on group cohesion. In 1998, Carron and colleagues, incorporating

the thrust of earlier definitions, defined cohesion as "a dynamic process which is reflected in the tendency for a group to stick together and remain united in the pursuit of its instrumental objectives and/or for the satisfaction of member affective needs." The definition underscores the dual focus of cohesion: task cohesion (i.e., instrumental in achieving group goals) and social cohesion (i.e., satisfying members' social needs). Groups can be characterized by task cohesion, social cohesion, or both. It is a dynamic process in that the attributes of the members and the group may change over time; thus, the nature and intensity of cohesion may also be in flux. For instance, a newly formed sport team may be marked by task cohesion in the beginning, and continued task interactions among members may lead to interpersonal liking and, thus, to social cohesion.

5.1. Multiple Dimensions of Cohesion

A cohesive group is characterized by (a) individual attraction (i.e., an individual values membership in the group) and (b) group integration (i.e., there are perceptions of positive and effective interactions among members). The team would be deemed more cohesive if all members were attracted to the group and if all members perceived the team to be integrated. By crossing the task versus social aspects of cohesion with individual attraction and group integration, Carron and colleagues derived the four-component model of cohesion (Fig. 1), consisting of (a) Attraction to the Group–Task (ATG-T), (b) Attraction to the Group–Social (ATG-S), (c) Group Integration–Task (GI-T), and (d) Group Integration–Social (GI-S).

5.2. The Group Environment Questionnaire

From a social cognitive perspective, Carron and colleagues argued that (a) members of a group selectively process the information they gain through experiences of the group's task and social properties and develop a set of beliefs about the group, (b) these member beliefs about the groups reflect the unity of the group, and (c) these social cognitions of members about the group reflecting its cohesiveness can be measured through a paper-and-pencil questionnaire. Accordingly, they developed a scale to measure the four components of the model. Their theory-driven Group Environment Questionnaire (GEQ) consists of 18 items to measure the four components of cohesion.

	Task Cohesion	Social Cohesion
Attraction to Group	Attraction to Group–Task	Attraction to Group–Social
Integration of Group	Group Integration–Task	Group Integration–Social

FIGURE 1 Multiple dimensions of cohesion. Adapted, by permission, from A. V. Carron, W. N. Widmeyer, L. R. Brawley, 1985, "The Development of an Instrument to Assess Cohesion in Sports Teams: The Group Environment Questionnaire," *Journal of Sport Psychology, 7*(3): Figure 2.

5.3. Correlates of Cohesion

The following is a summary of Carron and Hausenblas's 1998 discussion of the research on the relationships of cohesion in sport teams with other relevant variables. Higher perceptions of cohesion were associated with members' tendency to undertake more of the team activities and to share the responsibility for negative outcomes. When members perceived that the team was cohesive, they tended to attend practices and games more regularly and to put in extra effort during practice sessions. By the same token, higher cohesiveness was related to reduced "social loafing." Members' cognitive anxiety was lowered on teams that were perceived to be cohesive.

When members perceived the group to be integrated over the task at hand (GI-T), they expressed higher levels of role clarity, role acceptance, and role performance. With increasing task cohesion, members tended to be defensive of the team and to severely self-handicap (i.e., identify external causes for future failures). Perception of cohesion was also related to higher performance by the group. In a 2002 meta-analysis of studies that investigated relationships between cohesion and performance in sport teams, Carron and colleagues found that the significant relationship varied from moderate to large. Contrary to findings relating to groups in general, both task cohesion and social cohesion were significantly related to performance of sport teams. These authors also found that the relationship between cohesion and performance was stronger in women's teams than in men's teams. Hence, they suggested

that team-building efforts should focus on cultivating task-oriented attitudes and understandings as well as interpersonal interactions and liking among members. Such a strategy would be particularly effective in the case of women's teams.

A factor that was found to enhance cohesion within a team is size of the group. Small size of the group facilitated higher levels of cohesion. It is noteworthy that sacrifice on the part of members enhanced task and social cohesion, and this in turn led to the conformity to group norms. When leaders emphasized cohesion within the team and executed appropriate actions, cohesion was fostered. Most of the sport teams are characterized by a hierarchical distinction between the leader (i.e., the coach or manager) and the members (i.e., the players). The two major streams of research on leadership in sports are based on the Multidimensional Model of Leadership and the Mediational Model of Leadership.

6. THE MULTIDIMENSIONAL MODEL OF LEADERSHIP

Chelladurai's Multidimensional Model of Leadership (Fig. 2) is a synthesis and extension of several theories from mainstream industrial/organizational psychology. In the model, group performance and member satisfaction are considered to be a function of the congruence among three states (or aspects) of leader behavior: required, preferred, and actual. The antecedents of these three states of leader behaviors are the characteristics of the situation, the leader, and the members.

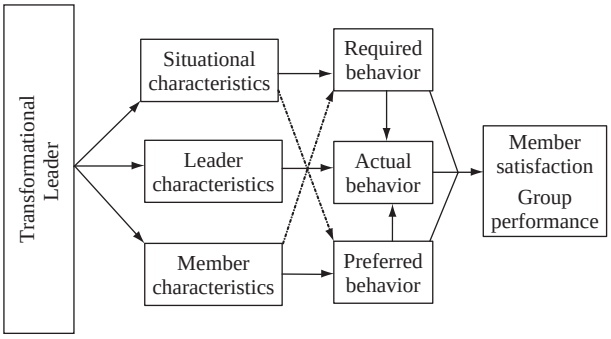


FIGURE 2 Multidimensional Model of Leadership. From *Handbook of Research on Sport Psychology*, by P. Chelladurai, Macmillan, ©1993 Macmillan. Reprinted by permission of the Gale Group.

6.1. States or Aspects of Leader Behavior

Required leader behavior refers to those behaviors shaped by the demands and constraints of the situational characteristics, that is, the parameters of the organization and its environment. The goals and formal organizational structure of the team and the larger system (e.g., professional team vs high school team), the group task and the associated technology (e.g., team sport vs individual sport), social norms, cultural values, and government regulations are some of the situational characteristics that impinge on leader behavior.

Members' preferences for specific leader behaviors, such as coaching and guidance, social support, and feedback, are largely a function of the individual characteristics (e.g., personality variables, age, experience, ability). In addition, the situational characteristics also affect member preferences. If there is an organizational expectation that a leader will behave in a specific manner, this expectation is shared by both the leader and members. Thus, both the leader and members are socialized into the same behavioral expectations and/or preferences in a given context.

A leader's actual behavior is seen as a function of his or her personal characteristics, including personality, ability, and experience. In addition, a leader is considerably influenced by the demands and constraints of the situation as well as by the preferences of the members. As noted previously, the major proposition of the multidimensional model is that the team's performance and member satisfaction are a function of the degree of congruence among the three states of leader behavior.

6.2. The Leadership Scale for Sports

In conjunction with proposing the multidimensional model, Chelladurai developed the Leadership Scale for Sports (LSS), which consists of 40 items representing the following five dimensions of leader behavior:

1. Training and instruction (emphasizing and facilitating hard and strenuous training; instructing in the skills, techniques, and tactics of the sport; clarifying the relationship between the members and their tasks; and structuring and coordinating the members' activities)
2. Democratic behavior (allowing greater athlete participation in decisions pertaining to group goals, practice methods, and game tactics and strategies)

3. Autocratic behavior (making decisions independently and stressing personal authority)
4. Social support (being concerned about the welfare of individual athletes, promoting a positive group atmosphere, and inducing warm interpersonal relations with members)
5. Positive feedback (reinforcing athletes by recognizing and rewarding good performance)

6.3. Correlates of Leadership

Studies have found that males preferred their coaches to be more training oriented and more autocratic yet more supportive than did female respondents. As for perceived behavior, male athletes perceived their coaches to place greater emphasis on training, instruction, and rewarding behavior and less on democratic behavior, whereas females perceived their coaches to be more democratic and socially supportive. However, there was a tendency at the top levels of competition for the coaches of male and female teams to be similar in their behaviors.

Athletes high on task motivation preferred more training and instruction, whereas those high on affiliation motivation and extrinsic motivation preferred more social support. Athletes high on cognitive structure preferred significantly more training and instruction and less autocratic behavior from the coach, whereas those high on impulsivity preferred more social support behavior from the coach.

With increasing experience, athletes preferred more positive feedback, autocratic behavior, and social support. Preference for social support increased progressively with maturity (i.e., from the high school level to the university level). In general, as athletes gained experience and/or ability, they tended to prefer their coaches to be more autocratic and socially supportive. That leads to the concept of the coach as the benevolent autocrat. Another general trend was for the coaches to rate themselves more positively on these behaviors (i.e., higher on training and instruction, democratic behavior, positive feedback, and social support) than the players did.

As for the situational variables, members of competitive sport teams preferred greater training and instruction, greater social support, less positive feedback, and less democratic behavior from their coaches than did members of recreational sport teams. As for task attributes, athletes involved in interdependent tasks (i.e., team sports) or variable tasks (i.e., open sports such as basketball) preferred more training and instruction and autocratic behavior than did athletes involved in independent tasks (i.e., individual sports) or nonvariable tasks (closed sports such as swimming). In general, increased task dependence and/or task variability seems to warrant the need for more training and instruction, autocratic behavior, social support, and positive feedback.

With regard to outcomes, athletes were satisfied with leadership to the extent that the coach emphasized (a) training and instruction that enhances the ability and coordinated effort by members that, in turn, contributes to task accomplishment, and (b) positive feedback that recognizes and rewards good performance. Performance as a consequence of leadership has been dealt with inadequately in the literature. The few studies indicate that higher perceptions of training, autocratic behavior, social support, and positive feedback were associated with success in athletics.

The relationships between leadership and performance were quite weak. Some scholars have used player perceptions as a measure of performance, and such perceptions may relate to performance and/or performance improvements. In contrast to performance, satisfaction has been studied more extensively. In 1998, Riemer and Chelladurai developed the Athlete Satisfaction Questionnaire (ASQ), consisting of 56 items to measure 15 facets of athlete satisfaction, including those relating to individual and team performance outcomes, individual and team social outcomes, individual and team task and social processes, and external agents.

7. THE MEDIATIONAL MODEL OF LEADERSHIP

The Mediational Model of Leadership (Fig. 3) was developed by Smoll and Smith with a focus on youth sport and

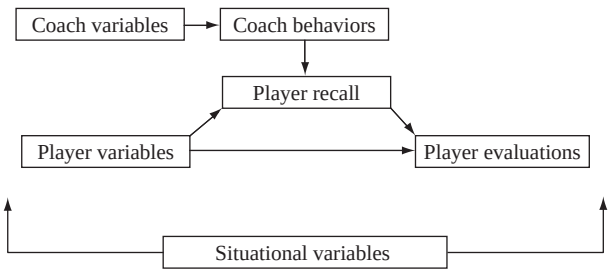


FIGURE 3 Mediational Model of Leadership. Reprinted with permission from *Journal of Applied Social Psychology*, Vol. 19, No. 18, pp. 1522–1551. ©V. H. Winston & Son, Inc., 360 South Ocean Boulevard, Palm Beach, FL 33480. All rights reserved.

emphasizes players' evaluative reactions as the outcome measures. The three critical elements of the model are coaches' behaviors, players' perceptions and recall, and players' evaluative reactions. Players' attitudes toward their coaches and their sport experience are mediated by their perceptions and recall of the coaches' behaviors. These three elements are affected by three sets of factors: coach individual difference variables, player individual difference variables, and situational factors.

7.1. Factors Influencing Leadership Dynamics

The coach individual difference factors include (a) coaches' goals, (b) coaches' behavioral intentions (i.e., the antecedents of actual behavior), (c) instrumentalities (i.e., a function of the perceived probability of an outcome and the value attached to that outcome), (d) the norms associated with the coaching role, (e) coaches' perceptions of players' motives, (f) self-monitoring (i.e., the extent to which coaches analyzes their own behaviors and their consequences), and (g) coaches' sex. Player individual difference variables include (a) the age that has been shown to affect players' perceptions and evaluations of their coaches, (b) players' sex, (c) players' perceptions of coaching norms, (d) the valence that players attach to various coaching behaviors, (e) players' achievement motive in the sporting context, (f) competitive trait anxiety, (g) general self-esteem (i.e., players' evaluation of general self-worth), and (h) athletic self-esteem (i.e., players' evaluations of themselves as athletes). The situational factors affect coaches' behaviors as well as players' perceptions of those behaviors and reactions to them. These factors include (a) nature of the sport, (b) level of competition (e.g., competitive sport vs recreational sport), (c) practice sessions versus game sessions, (d) previous success/failure record of the team, (e) current game/practice developments (e.g., whether the team is losing or winning), and (f) intrateam attraction.

7.2. The Coaching Behavior Assessment System

Smoll and Smith employed the observational method to assess the actual leader behavior and employed the paper-and-pencil method to assess players' perceptions and recall of that leader behavior as well as coaches' perceptions of their own behaviors. These authors' Coaching Behavior Assessment System (CBAS) was developed over several years by observing and recording the behaviors of youth soccer basketball and baseball coaches, analyzing these recordings, and categorizing the relevant leader behaviors. The 12 behavioral dimensions measured by the CBAS are broadly classified into reactive and spontaneous behaviors. Reactive behaviors are coaches' responses to (a) desirable performance or effort, (b) mistakes and errors, and (c) misbehaviors. Spontaneous behaviors may be (a) game related or (b) game irrelevant.

The training of the observers has been quite extensive, including a programmed learning module, group instruction in use of the coding system, written tests, practice, and reliability checks. Such training resulted in the coders agreeing with the experts more than 90% of the time. The second coding of the same coaching behaviors after a lapse of 1 week was also consistent to the tune of 87.5% to 100%, with a mean of 96.4%.

To assess players' perceptions and recall of coaches' behaviors, players were provided with a description and an example of each of the 12 behavioral dimensions of the CBAS and were asked to indicate the extent to which their coaches had engaged in each of the behaviors. A similar format was followed in assessing coaches' perceptions of their own behaviors. The relationships among the three measures of coaching behavior—observation, coaches' self-reports, and players' perceptions and recall—were not strong. More specifically, coaches' self-reports of their behaviors did not bear much resemblance to the assessments made by either observers or players. Therefore, the use of coaches' self-reports of their behaviors should be viewed with extreme caution. Player evaluative reactions were scored with a 10-item scale measuring attraction of the coaches and attraction toward teammates.

7.3. Correlates of Leadership

In general, leaders' supportiveness (i.e., reinforcement and mistake-contingent encouragement) and instructiveness (i.e., general technical instruction and mistake-contingent instruction) were significantly and positively related to players' attitudes toward their coaches, the sport, and teammates, whereas punishment and punitive technical instruction had the opposite effect. Also, coaching behaviors affected players' liking for the sport after the season, their perceptions of team solidarity, their evaluation of their coaches, and their

self-esteem. Although general technical instruction was found to be more potent than the other categories, positive reinforcement had minimal effects.

Player self-esteem has been used both as (a) a moderator of the relationship between coaches' behaviors and player attitudes and as (b) an outcome measure. Children with low self-esteem tended to respond most positively to a supportive or an instructive coach–player relationship and tended to react most negatively to coaches who were not supportive or instructive. It was also found that coaching behaviors had far less impact on children with high self-esteem. As for self-esteem as an outcome, children playing for trained coaches exhibited higher levels of self-esteem than they had a year before. Such increases were not found among children playing for nontrained coaches. Coaching behaviors during practices were also shown to be significantly associated with changes in players' perceptions of their physical and cognitive competence as well as success expectancy.

The mediational model has been the foundation for training coaches to improve their behaviors, assess their behaviors, and measure the effects of changes in coaching behaviors on players' evaluative reactions. Research has shown that the trained group of coaches differed from the control group in both observed and perceived behaviors. That is, the former group engaged in behaviors consistent with the training program. Furthermore, the players of the trained coaches evaluated their coaches more positively and expressed higher levels of intrateam attraction than did the players of the control group of coaches. However, the two groups did not differ in performance as measured by win/loss records.

7.4. Commentary on Leadership Studies in Sport

Some hold that the Multidimensional Model of Leadership is meant for adult sport and that the Mediational Model of Leadership is meant for youth sport because these models were developed in the context of adult sport and youth sport, respectively, and have been employed predominantly in adult sport and youth sport, respectively. But there is nothing in either of the models to prohibit or inhibit its use in either level of sport. Although there are some conceptual differences between the two models, each of them is applicable to either level.

The distinguishing feature of the research carried out by Smoll and Smith is the manner in which leader behavior is measured, that is, the observational method of the CBAS versus the paper-and-pencil LSS. However, a comparison of leadership behavior categories of the two systems shows that they are, to a large extent, similar to each other in content, although the labels are different. For example, the reinforcement and nonreinforcement categories of the CBAS resemble the LSS dimension of positive feedback; the mistake-contingent technical instruction, general technical instruction, and organization are subsumed by the LSS dimension of training and instruction; and general communication is similar to social support in the LSS. Despite these similarities, the CBAS is distinct in two significant respects. First, the larger number of categories in the CBAS permits a more thorough analysis of leadership in youth sport. Apart from this advantage, the CBAS also measures leader behavior in more specific situations. For example, although "encouragement" in the CBAS is subsumed by the "positive feedback" in the LSS, encouragement is divided into reactive mistake-contingent behavior and spontaneous general encouragement.

In 2002, Horn employed a working model as the framework for her review of coaching effectiveness. While incorporating significant elements of both the multidimensional and mediational models of leadership, she also introduced the construct of "coaches' expectancies, values, beliefs, and goals" as a factor mediating the effects of the "sociocultural context," "organizational climate," and "coaches' personal characteristics" on "coaches' behavior." Furthermore, she referred to self-efficacy, motivational style (i.e., intrinsic or extrinsic motivation), self-reflectiveness, critical thinking aptitudes, locus of control, and trait anxiety as possible predictors of coaches' expectancies, values, beliefs, and goals. Coaches' expectations referred to coaches developing expectations for the performance of their teams and their members that affects coaches' subsequent behaviors. Beliefs were narrowed down to stereotyped beliefs about gender, race and ethnicity, and sexual orientation.

8. MOTIVATIONAL CLIMATE

An emergent thrust in the study of group dynamics pertains to the motivational climate within sport groups created and fostered by coaches. The motivational climate is characterized as (a) task involving (i.e., mastery orientation) or (b) ego involving (i.e., performance orientation). In the climate of mastery orientation, hard work, trying hard, skill improvement, and contributions by all members are emphasized. In

contrast, comparisons with others and performing better than others are promoted in the performance-oriented climate. Research has shown that, in general, the mastery-oriented (task-involving) climate enhances athletes' level of enjoyment, satisfaction, and interest in the sport, whereas the performance-oriented (ego-involving) climate tends to increase tension, pressure, and performance anxiety. Insofar as coaches have considerable influence and control over the climate in which their teams operate, it could be speculated that coaches' own goal orientation would be reflected in their leadership behaviors, which in turn would create either the mastery-oriented climate or the performance-oriented climate. It is also possible that the organization itself may foster the performance orientation that impinges on both coaches and athletes. For example, coaches of Division I schools in the National Collegiate Athletic Association (NCAA) are likely to exhibit those behaviors that would foster a performance-oriented climate, and athletes are likely to endorse that orientation and flourish in it. The relative influences of the organizational goals and coaches' own goal orientation on the climate experienced by team members are yet to be determined. Finally, the idea that pursuit of excellence would be greatly facilitated if the climate reflected both task orientation and ego orientation is also worthy of investigation.

See Also the Following Articles

Goal Setting and Achievement Motivation in Sport ■ Leadership and Culture ■ Power, Authority, and Leadership

Further Reading

Carron, A. V., Brawley, L. R., & Widmeyer, W. N. (1998). The measurement of cohesiveness in sport groups. In J. L. Duda (Ed.), *Advancements in sport and exercise psychology measurement* (pp. 213–226). Morgantown, WV: Fitness Information Technology.

Carron, A. V., Colman, M. M., Wheeler, J., & Stevens, D. (2002). Cohesion and performance: A meta-analysis. *Journal of Sport and Exercise Psychology*, *24*, 168–188.

Carron, A. V., & Hausenblas, H. A. (1998). *Group dynamics in sport*. Morgantown, WV: Fitness Information Technology.

Chelladurai, P. (1993). Leadership. In R. N. Singer, M. Murphy, & K. Tennant (Eds.), *The handbook on research in sport psychology* (pp. 647–671). New York: Macmillan.

Chelladurai, P., & Saleh, S. D. (1980). Dimensions of leader behavior in sports: Development of a leadership scale. *Journal of Sport Psychology*, *2*(1), 34–45.

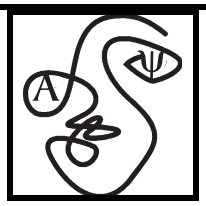
Duda, J. L., & Hall, H. (2001). Achievement goal theory in sport: Recent extensions and future directions. In R. N. Singer, H. A. Hausenblas, & C. M. Janelle(Eds.), *Hand book of sport psychology* (2nd ed., pp. 417–443). New York: John Wiley.

Horn, T. S. (2002). Coaching effectiveness in the sport domain. In T. S. Horn (Eds.), *Advances in sport psychology* (2nd ed., pp. 309–354). Champaign, IL: Human Kinetics.

Riemer, H. A., & Chelladurai, P. (1998). Development of the Athlete Satisfaction Questionnaire (ASQ). *Journal of Sport and Exercise Psychology*, *20*, 127–156.

Smoll, F. L., & Smith, R. E. (1989). Leadership behaviors in sport: A theoretical model and research paradigm. *Journal of Applied Social Psychology*, *19*, 1522–1551.

Widmeyer, W. N., Brawley, L. R., & Carron, A. V. (2002). Group dynamics in sport. In T. S. Horn (Ed.), *Advances in sport psychology* (2nd ed., pp. 285–308). Champaign, IL: Human Kinetics.



Groups, Productivity within

Athena Xenikou
Aristotle University of Thessaloniki, Thessaloniki, Greece

Alexandra Hantzi
Panteion University, Athens, Greece

1. Introduction
2. Social Facilitation/Inhibition
3. Steiner's Theory of Group Productivity
4. Social Loafing
5. Motivation Gains in Groups
6. Implications for Groups at Work
Further Reading

Research on group performance mainly examines the conditions under which motivation losses and gains are produced when people work in groups. Motivation losses concern the reduction in individual effort to perform a task due to the fact that one works as part of a group. On the other hand, it has been suggested that motivation gains resulting from group settings need to be more thoroughly investigated.

GLOSSARY

free-rider effect The enjoyment of the benefits resulting from successful group performance without contributing to it.
production blocking Production blocking in brainstorming groups arises from the constraint that only one group member can talk at a time.
social compensation A group member compensating for the low performance of other group members for the desired collective goal to be accomplished.
social facilitation/inhibition effects The facilitation or impairment of individual task performance as a function of the presence of others either as an audience or as coactors.
social loafing The decrease in individual effort when one performs as a member of a group in comparison to performing alone.
sucker effect The expectation that others withhold efforts in performance groups, which can lead an individual reducing his or her contribution to prevent being a victim of free-riding.
transactive memory A knowledge shared by group members about who is good at what.

1. INTRODUCTION

Most people have at some point wondered if it is more productive to work as part of a team or alone. This is a simple but challenging question, which social psychology has faced since it appeared as a discipline. Researchers have worked hard for decades to answer whether group performance is superior to individual performance. In general, the findings have shown that groups are more effective than separate individuals because groups learn faster, produce a greater quantity and quality of work, and make fewer errors.

However, groups do not accomplish their full potential since their performance is not as high as the sum-total of the maximum individual performance of each group member. It has been suggested that coordination losses and motivation losses are responsible for groups performing less efficiently than expected. Researchers have also identified conditions under which motivation gains result out of working as part of a group. Social compensation refers to the increased effort of a group

©2004 Elsevier Inc.
All rights reserved.

member in order to compensate for the insufficient contribution of another member when the group product is perceived as valuable. Moreover, there is evidence that complex and self-involving tasks carried out by groups that are psychologically meaningful to their members do not lead to productivity losses but gains.

2. SOCIAL FACILITATION/ INHIBITION

The most elementary issue concerning the impact of social factors on individual task performance is whether the mere presence of others influences individual performance. Social facilitation and inhibition (SFI) refer to the positive and negative effects that audiences (as well as coactors) may have on individual task performance. Interest in how the presence of others affects individual performance dates back to 1898 when Tripplett conducted one of the first experiments in the field of social psychology. Tripplett had children spin fishing reels either alone or in coactive pairs and found that participants performed better coactively. Following Tripplett's work, many researchers replicated the finding that the presence of others facilitates individual performance, whereas others demonstrated the social inhibition of individual performance under the same conditions.

In 1965, these apparently inconsistent results were explained by Robert Zajonc, who proposed that the presence of others facilitates individual performance in the case of simple tasks, whereas it impairs performance at complex tasks. Zajonc put forward a drive theory of social facilitation suggesting that the mere presence of others increases generalized drive or physiological arousal. Arousal, in turn, leads to the generation of dominant responses—that is, responses that are well learned, habitual, or innate. Simple tasks are successfully performed when dominant responses are generated, whereas complex tasks being more demanding in nature cannot be accomplished by the emission of dominant responses.

A number of other theoretical accounts of how task complexity affects individual performance at social settings have been provided. Zajonc had such a great impact that some of these theoretical accounts did not challenge the drive approach of his theory and presented factors considered as determinants of drive/arousal. For example, (i) people's concern with evaluation and (ii) distraction/conflict caused by attending both to the task at hand and to other people were factors proposed by different theoretical approaches to account for elevated arousal at social settings. Recent theories have challenged the biologically based explanation offered by Zajonc's theory and have focused on cognitive and attentional factors as determinants of SFI effects.

There is a strong empirical basis suggesting that task complexity is related to the manifestation of SFI effects as proposed by Zajonc. However, it has been found that physiological arousal is not increased by the presence of others irrespective of task complexity. Complex tasks do lead to an elevated drive when others are present, whereas simple tasks do not seem to provoke increased levels of physiological arousal. Finally, recent research, in accordance with an attentional view of social facilitation, has shown that the presence of unpredictable others facilitated performance at a complex task by the emission of a nondominant response. In summary, Zajonc's proposition that task complexity plays an important role in explaining SFI effects has been strongly supported by the findings, but the exact mechanism(s) leading to facilitation and impairment at simple and complex tasks, respectively, has not been identified.

3. STEINER'S THEORY OF GROUP PRODUCTIVITY

Although the literature on social facilitation suggests that under certain conditions one feature of group settings (i.e., the presence of others) enhances individual performance, most of the work on group performance argues that when people work together in groups they do not perform at the maximum level of their potential. In 1972, Steiner proposed an influential analysis of group productivity, which focuses on the notion of a group's potential productivity, the maximum possible performance level that the group can achieve. Potential productivity is determined by group members' resources and task demands. Member resources concern the capabilities, knowledge, and skills of group members that can be useful at completing the task at hand. Task demands mainly refer to the criterion that is used to estimate a group's performance level. The criterion of performance at a weight-lifting task is the total weight lifted, whereas in the case of a brainstorming task it is the number of novel ideas generated by the group.

Steiner's model assumes that a group is not in a position to achieve its potential productivity due to process loss (faulty group process), and that a group's actual productivity is potential productivity minus losses due to faulty process. Process loss refers to coordination as

well as motivation losses. When performing in groups it is necessary for group members to coordinate their actions to achieve the desired group goal. For example, in a rope-pulling task group members have to synchronize their actions to achieve their full potential. This need for coordination inevitably leads to loss in productivity since group members cannot pull at exactly the same time or direction. Regarding motivation losses, individuals are considered to exert less effort when working as part of a team than when working alone.

Steiner also provided a task classification or taxonomy (Table I), which aimed at further assisting researchers in the investigation of group productivity. According to Steiner's classification, there are three task characteristics that can be used to categorize different tasks. The first task feature is whether a task can be subdivided to separate parts performed by different people. The tasks that cannot be subdivided are unitary, whereas the tasks that can be subdivided are divisible. An example of a unitary task is finding a solution to a mathematical problem. On the other hand, a task that certainly (and hopefully) can be divisible among roommates is cleaning their apartment. The second task feature that Steiner used to classify tasks is the performance criterion mentioned earlier, leading to a distinction between maximizing and optimizing tasks. When a task performance criterion concerns the quantity or speed of production, the task is considered to be a maximizing task; the rope-pulling task is a good example of a maximizing task. On the other hand, optimizing tasks put a demand for a correct or optimal solution (e.g., a mathematical problem). Finally, the way that group members' inputs are related to the group product leads to four different types of tasks: disjunctive, conjunctive, additive, and discretionary tasks. In a disjunctive task, members make separate contributions and the best contribution must be chosen (e.g., a reasoning problem). In conjunctive tasks, group performance is determined by the contribution of the least capable member (e.g., climbing a mountain while members are tied to one another). In additive tasks, the individual contributions are simply added together (e.g., a rope-pulling task). Finally, in discretionary tasks, individual inputs are combined in any way the group chooses (e.g., roommates cleaning their apartment).

TABLE I
Steiner's Task Classification[a]

Task characteristics	*Task taxonomy*
Can a task be subdivided to separate parts performed by different people?	Yes → *Divisible tasks* No → *Unitary tasks*
Does the task performance criterion concern quantity/speed of production or a correct/optimal solution?	The performance criterion concerns Quantity/speed of production → *Maximizing tasks* Correct/optimal solution → *Optimizing tasks*
How are group members' inputs related to group product?	Members make separate contributions and the best contribution must be chosen → *Disjunctive tasks* Group performance is determined by the contribution of the least capable member → *Conjunctive tasks* The individual contributions are simply added together → *Additive tasks*

[a]Adapted from Steiner (1972).

Steiner's model has been criticized on the basis that it explicitly makes the assumption that groups cannot reach their full potential. It has been argued that groups might be in a position not only to achieve performance at the highest level of their potential but also to overperform it. Steiner's theory has put emphasis on coordination and motivation losses that influence group performance, while discounting the possibility that groups may exhibit motivation gains.

4. SOCIAL LOAFING

Productivity loss in group settings has been defined in terms of both coordination and motivation losses. The first experiments on group productivity were conducted during the last decades of the 19th century by Ringelmann. In these experiments, individuals were asked to pull a rope either alone or in groups of various sizes. The results showed that when individuals worked in groups, the collective output was less that the sum total of the individual performance measured when individuals pulled the rope alone. This finding can be explained as the result of coordination loss. However, subsequent research showed that this finding

is caused by the joint effect of coordination problems and reduced effort exerted by group members.

Regarding motivation losses, there is a difference between the effort that individuals exert when working collectively and working coactively (or individually). When people work collectively, their individual contributions are combined to form a group product, whereas when they work coactively their individual inputs are not combined to a single product. Individuals' tendency to produce lower effort when working collectively compared to working coactively (or individually) was called social loafing by Latané and colleagues in 1979.

Social loafing has been observed in a wide variety of tasks, including physical tasks (e.g., rope pulling, shouting, and clapping), cognitive tasks (e.g., idea-generation tasks), perceptual tasks (e.g., maze performance), and evaluative tasks (e.g., evaluation of poems and editorials), suggesting that social loafing is a robust phenomenon. As a general phenomenon not restricted to physical tasks, it poses the important practical question of why and when less effort is exerted in group settings.

A number of theoretical explanations have been proposed to explain why social loafing takes place in groups. It has been suggested that social loafing occurs because (i) the pressure people feel as group members to accomplish the task at hand is less intense than the pressure one feels when working alone, (ii) the presence of other group members is drive reducing because others share the responsibility for carrying out the task, (iii) there is reduced identifiability or evaluation potential in collective action compared to individual action, and (iv) group members feel their contributions to the group product not to be essential but dispensable.

Apart from social loafing, two other effects have been reported, namely, the free-rider effect and the sucker effect. The free-rider effect concerns the enjoyment of the benefits resulting from the successful group performance without contributing to it. The free-rider effect might be seen as the result of a person's conscious decision to indulge in social loafing. On the other hand, members of a group who work hard to accomplish the group goal run the risk that others may free ride on their efforts and become suckers. The expectation that others withhold efforts in performance groups can lead to the individual reducing his or her contribution to the group in order to prevent being a victim of free-riding (the sucker effect).

The literature that examines group conditions leading to social loafing has also attempted to propose various solutions to remedy motivation loss in group settings. On a practical level, the following can reduce or overcome social loafing in real-world contexts: (i) identification and evaluation of individual inputs, (ii) meaningful and self-involving tasks, (iii) close ties among group members, (iv) small rather than large groups, (v) uniqueness of individual contributions, and (vi) feedback on individual or group performance.

5. MOTIVATION GAINS IN GROUPS

Even though research on group performance has concentrated on productivity loss in group settings due to reduced motivation, there are findings showing that motivation gains can also result from performing within groups. In 1991, Williams and Karau demonstrated conditions under which people may actually increase their efforts when working collectively compared to when working individually (coactively). They call this effect social compensation. Social compensation refers to a group member compensating for the low performance of other members in order to accomplish the desired group goal.

One factor that might produce social compensation is the expectation that other group members are performing insufficiently. The perception of inadequate coworker inputs may be derived from a general lack of trust in the reliability of others to perform well when their contributions are pooled with those of others or/and direct knowledge of coworkers' insufficient efforts or inabilities. The second factor that may be necessary to produce social compensation is that the group product is in some way important to the individual. If the group output is not meaningful and important to group members, then one does not expect social compensation to occur.

Social compensation is not expected to occur in all situations in which individuals believe their coworkers are contributing less than their fair share. Besides group product meaningfulness previously mentioned, other factors are important for compensation to take place. Social compensation is more likely to occur (i) when the individual has to remain in the group working collectively, (ii) at earlier stages of the collective effort, and (iii) when the group size is relatively small. Finally, people may be motivated to compensate in some cases for their own prior underperformance rather than the underperformance of other group members.

According to social compensation, when group members of unequal strength work on an important task, the most competent member will compensate for the insufficient contribution of the other group

member(s). There are, however, findings showing that under specific conditions the weaker member of the group tries hard to match the performance level of the stronger member (the Köhler effect). One necessary condition for this effect to occur is that the group product must be meaningful to its members. The second condition is that the discrepancy in ability between the group members is of a moderate degree. A high discrepancy in individual performance is not expected to lead to more effort exerted by the weakest members because people believe that there is a limit to the discrepancy in ability that can be compensated for through increased effort.

The social loafing paradigm has been criticized on the basis that little effort is expended in making the participants feel that they are part of a group. In the paradigm, a number of individuals combine their efforts for a group output without interacting with each other, competing with other groups, or expecting any form of future interaction. Even though the findings of research on social loafing are discussed in terms of group membership, it seems that participants mainly act on the basis of their personal identity rather than a group-based sense of self. Motivation losses described by research on social loafing might be the result not of faulty group processes but of individuals acting as persons rather than group members. It has been demonstrated that when social identity is made salient there are motivation gains leading to the accomplishment of the desired goal as defined by the group rather than some external source of influence.

6. IMPLICATIONS FOR GROUPS AT WORK

Despite the fact that most research on group performance has emphasized the inefficiency of groups compared to individuals, groups feature in most essential aspects of everyday life. Their central role in real-world contexts might have to do with the fact that many tasks cannot be accomplished by individuals. If one considers how many people with different skills are involved in making a car or building a house, then it becomes clear that getting rid of groups cannot be a practical solution to increase performance.

Regarding groups at work, how can we use the accumulated knowledge on group performance to understand how groups function and to propose ways on a practical level for groups to become more effective?

In organizational settings, brainstorming groups are used in order for the organization to stimulate employees to think creatively. In brainstorming groups, members are asked to generate as many ideas as possible, not to criticize others' ideas, and to build on ideas presented by other group members. Research by Diehl and Stroebe in the late 1980s showed that brainstorming groups produce half the ideas produced by the same number of individuals who brainstorm alone (nominal groups). This productivity loss in brainstorming groups has been discussed as a coordination loss resulting from mutual production blocking. Mutual production blocking arises from the constraint that only one group member can talk at a time in a brainstorming group.

Although there is ample evidence showing brainstorming groups to be less effective than nominal groups, organizations of all kinds make wide use of brainstorming groups. A reason that might explain organizational members' confidence in brainstorming groups is that members of such groups are more satisfied with their performance in comparison to individuals who brainstorm alone. This discrepancy between satisfaction and productivity can be accounted for by the greater ease of idea generation in brainstorming groups compared to nominal groups. Since brainstorming is widely used at work, some advice needs to be given on the conditions that make brainstorming groups more effective. First, the size of the group should be kept as small as possible, not exceeding four members. Second, because time pressure negatively affects the productivity of brainstorming groups, groups should decide when it is the best time to stop. Finally, it is advisable that group members are heterogeneous in their knowledge about the brainstorming topic so that members can have a stimulating impact on each other.

Moreover, our knowledge on group productivity can be applied to training at work in order to improve organizational effectiveness. Interventions designed to increase productivity through learning, such as training programs, were found to have a strong positive effect. Regarding group training programs, it has been suggested that group learning is associated with a transactive memory system—that is, a knowledge shared by group members on who is good at what. When members of a work group are trained together, they develop complementary specializations that enable the group to share knowledge and skills. In this way, the group as a whole makes use of social compensation—that is, individual members compensate for each other's inadequacy. However, a practical problem related to group training at work involves the

disruption that employee turnover has on the smooth operation of the work group.

The business world has acknowledged the importance of one's ability to work within groups as a key factor to success. In this sense, groups are seen as capable of accomplishing more than the sum of their individual parts. However, the bulk of research on group performance suggests that, more often than not, groups are not as efficient as individuals.

See Also the Following Articles

Conflict within Organizations ■ Groupthink ■ Intergroup Relations and Culture ■ Interpersonal Conflict

Further Reading

Baron, R. S., & Kerr, N. L. (2003). *Group process, group decision, group action*. Buckingham, UK: Open University Press.

Brown, R. (2000). *Group processes: Dynamics within and between groups*. Oxford, UK: Blackwell.

Diehl, M., & Stroebe, W. (1987). Productivity loss in brainstorming groups: Toward the solution of a riddle. *Journal of Personality and Social Psychology*, *53*, 497–509.

Haslam, A. (2001). *Psychology in organizations: The social identity approach*. London: Sage.

Latané, B., Williams, K., & Harkins, S. (1979). Many hands make light work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology*, *37*, 822–832.

Ringelmann, M. (1913). Recherches sur les moteurs animes: Travail de l'homme [Research on animate sources of power: The work of man]. *Annales de l' Institut National Agronomique*, *12*, 1–40.

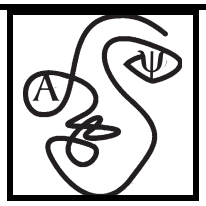
Steiner, I. D. (1972). *Group process and productivity*. New York: Academic Press.

Stroebe, W., Diehl, M., & Abakoumkin, G. (1996). Social compensation and the Kfhler effect: Toward a theoretical explanation of motivation gains in group productivity. In E. Witte, & J. H. Davis (Eds.), *Understanding group behavior: Small group processes and interpersonal relations* (pp. 37–65). Mahwah, NJ: Lawrence Erlbaum Associates.

Tripplett, N. (1898). The dynamogenetic factors in pacemaking and competition. *American Journal of Psychology*, *9*, 507–533.

Williams, K. D., & Karau, S. J. (1991). Social loafing and social compensation: The effects of expectations of coworker performance. *Journal of Personality and Social Psychology*, *61*, 570–581.

Zajonc, R. B. (1965). Social facilitation. *Science*, *149, 269–274.*



Groupthink

Ramon J. Aldag
University of Wisconsin, Madison, Wisconsin, USA

Sally Riggs Fuller
University of South Florida, Tampa, Florida, USA

1. The Groupthink Model Defined and Explained
2. History and Impact of the Groupthink Phenomenon
3. Evidence Regarding the Groupthink Phenomenon
4. Bases for Groupthink's Appeal
5. Implications and Suggestions for Future Research and Theory
6. Prescriptions for Effective Group Problem Solving
Further Reading

GLOSSARY

availability heuristic Assessment of the frequency of a class or the probability of an event based on the ease with which instances or occurrences can be brought to mind.

defensive avoidance A condition in which an individual fails to engage in vigilant search and appraisal and instead adopts a dominant mode of coping that involves avoidance of cues that stimulate anxiety or other painful feelings.

groupthink A condition in which highly cohesive groups in "hot" decision situations display excessive levels of concurrence seeking that suppress critical inquiry and result in faulty decision making.

"hot" cognitions Thought processes regarding vital, affect-laden issues (i.e., in "hot" decision situations).

illusory correlation The tendency to see relations between variables where none exists.

mindguard In groupthink, a self-appointed individual who acts to shield a group from adverse information.

Pygmalion effect Self-fulfilling prophecy; it is the tendency to create something in the image that one has of it.

strong version of groupthink The view that groupthink characteristics cluster due to their common ties to specific antecedent conditions and are linked in a causal chain from those antecedent conditions, to symptoms, to defects, to outcomes.

weak version of groupthink The view that groupthink may be evidenced by the presence of some subset of the characteristics proposed by Janis and that the causal ordering that he posited may be suggestive rather than necessary.

Groupthink is a condition in which highly cohesive groups in "hot" decision situations display excessive levels of concurrence seeking that suppress critical inquiry and result in faulty decision making. In the groupthink process, antecedent conditions are seen as resulting in a variety of symptoms of groupthink that lead to defective decision making and, in turn, to flawed decision outcomes. This article addresses the nature and history of the phenomenon, offers evidence of its validity and explanations for its widespread acceptance, and provides implications and suggestions.

1. THE GROUPTHINK MODEL DEFINED AND EXPLAINED

Janis proposed that highly cohesive groups are likely to suffer from groupthink, a strong concurrence-seeking tendency that suppresses critical inquiry and results in faulty decision-making processes and flawed outcomes. Janis discussed major historical fiascoes such as the lack of preparedness for the Japanese attack on Pearl Harbor,

©2004 Elsevier Inc.
All rights reserved.

the escalation of war in Korea, the failed U.S.-sponsored landing of anti-Castro rebels in the Bay of Pigs, and escalation of U.S. involvement in the war in Vietnam. First presented in a 1971 issue of *Psychology Today,* this groupthink phenomenon quickly gained remarkably broad and firm acceptance, dominating the literature on group decision making for decades.

At the individual level, Janis and Mann argued that thinking about vital, affect-laden issues generally results in "hot" cognitions, in contrast to the "cold" cognitions of routine problem solving, and that such hot cognitions can result in errors in scanning of alternatives. Such situations induce stress that results in defensive avoidance, characterized by lack of vigilant search, distortion of the meanings of warning messages, selective inattention and forgetting, and rationalizing. Groupthink is seen as the group analog of defensive avoidance or "a collective pattern of defensive avoidance." In 1972, Janis wrote that he used the term "groupthink" to refer to a mode of thinking engaged in by people in cohesive in-groups, with members' strivings for unanimity overriding their motivation to submit alternative courses of action to realistic appraisal. He further stated, "Groupthink refers to a deterioration of mental efficiency, reality testing, and moral judgment that results from in-group pressures."

Janis chose the term "groupthink" due to its frankly Orwellian connotation, similar to "doublethink" and "crimethink." In 1982, Janis wrote that "the invidiousness is intentional." This view of groupthink as an undesirable phenomenon continued to be evidenced in Janis's more recent work.

Janis explicitly defined the context in which groupthink is thought to occur and divided these antecedent conditions into three categories. First, a moderate to high level of group cohesion is a necessary, but not sufficient, condition for groupthink. In addition, structural faults and a provocative situational context are secondary antecedents that define situations in which groupthink is likely to occur. In the structural fault category are insulation of the group, lack of impartial leadership, lack of norms requiring methodical procedures, and homogeneity of group members' social backgrounds and ideologies. The provocative situational context antecedents focus on the role of stress as a situational factor. The first stress factor is characterized as external threats of losses combined with a low hope of finding a better solution than that of the leader. The internal stress antecedent stems from temporary low self-esteem attributable to members' recent failures, perceptions that the task is too difficult to accomplish, and the perception that there is no alternative that is morally correct.

According to Janis, the antecedents lead to several proposed symptoms of groupthink, including an illusion of invulnerability, rationalization to discount warnings and other negative feedback, a belief in the inherent morality of the group, stereotyped views of members of opposing groups, pressure on dissenters, self-censorship, illusion of unanimity, and self-appointed "mindguards" acting to shield the group from adverse information.

According to Janis, groupthink results in a number of consequences that interfere with effective group decision making. For instance, the group limits its discussion to only a few alternatives. After a course of action is initially selected, group members ignore new information concerning its risks and drawbacks. At the same time, they avoid information concerning the benefits of rejected alternatives. Members make little attempt to use experts to obtain more precise information. In addition, because they are so confident that things will turn out well, members fail to consider what may go wrong and, as such, do not develop contingency plans. These "defects" are seen as leading to impaired performance and other undesirable outcomes.

Janis suggested methods to prevent or minimize the supposedly dysfunctional consequences of groupthink. These "remedies" included the following:

- Group leaders should encourage all group members to air their doubts and objections.
- Group leaders should adopt an impartial stance rather than initially stating their preferences.
- Group members should be encouraged to discuss the group's deliberations with trusted associates and report their reactions back to the group.
- Outside experts should be invited to meetings and encouraged to challenge group members' views.
- When a competitor is involved, time should be devoted to assessment of warning signals from the competitor and of alternative scenarios of the competitor's intentions.
- When considering alternatives, the group should split into subgroups to meet separately from time to time.
- The group should hold a "second chance" meeting after a preliminary consensus is reached on a preferred alternative.
- The group should consider using dissonance-inducing group processes.

2. HISTORY AND IMPACT OF THE GROUPTHINK PHENOMENON

More than 30 years after its conception, the groupthink phenomenon retains a remarkably strong intuitive appeal and acceptance. A recent Google search revealed more than 42,700 groupthink "hits." Groupthink is presented as received doctrine in sources ranging from *Educational Gerontology* to the *Utne Reader,* from the *New Criterion* to *Vogue,* and is offered as the cause for everything from flawed jury deliberations, to problems of the Washington Redskins, to the U.S. decision to invade Iraq. The *New York Times* recently reported that groupthink is three decades old and still going strong, invoking the phenomenon as a potential explanation for the space shuttle *Columbia* disaster. In addition to its firmly entrenched status with practitioners in a wide variety of areas, groupthink continues to be featured in popular textbooks and to be the subject of theory and research.

3. EVIDENCE REGARDING THE GROUPTHINK PHENOMENON

Most support for groupthink has come from retrospective case studies that have focused on decision fiascoes rather than comparing the decision-making processes associated with good decisions with those associated with bad decisions.

Experimental studies of groupthink have considered only a small portion of the model, often without a cohesive group and in situations inconsistent with Janis's proposed antecedents. Furthermore, they have relied nearly exclusively on student samples dealing with hypothetical or simulated decisions, with potential resultant problems for external validity. In the laboratory, many real-world group characteristics, including ongoing power relationships and political maneuverings, have been necessarily ignored. Empirical studies of groupthink have tended to use short-term decision quality as the sole outcome measure and have contained serious threats to internal validity. Interestingly, Janis apparently discounted these experiments. For instance, in his 1989 book, he cited none of the experimental studies of groupthink, with his discussion of groupthink referring only to case studies and anecdotal pieces.

Another important aspect of prior research concerns the situations chosen for study. The groupthink conceptualization is restrictive, primarily due to the model's antecedents and the international policy context for which groupthink was developed. The narrow focus of the model's antecedents is not surprising given the fact that Janis extracted them directly from major policy decision fiascoes. Early case evidence provided in support of groupthink was drawn largely from these and similar hot decision situations. The ability to generalize from such situations is questionable. For instance, the extremely high group cohesiveness and clearly unified goals in many major military policy decisions, or in groups facing natural disasters, are unlikely to be encountered often in other settings. Nevertheless, laboratory and survey research on groupthink has typically used settings considerably different from Janis's conceptualization.

Strong and weak interpretations of groupthink have emerged. The strong interpretation, commonly cited, suggests that groupthink is an integrated set of characteristics with deterministic linkages. That is, groupthink characteristics cluster due to their common ties to specific antecedent conditions and are linked in a causal chain from those antecedent conditions, to symptoms, to defects, to outcomes. The weak interpretation is sometimes presented in response to the failure of empirical examinations to provide results that are wholly consistent with groupthink. It suggests that groupthink may be evidenced by the presence of some subset of these characteristics and that the causal ordering posited by Janis may be suggestive rather than necessary.

When the weak interpretation is presented, it is crucial to ask which characteristics, if any, are unique to groupthink. That is, some groupthink characteristics, such as belief in the correctness of a decision and pressures to achieve consensus, are common to many group processes and, in fact, may be beneficial. For instance, convergence to the majority viewpoint may be desirable if the majority viewpoint is correct. Other characteristics, such as mindguards and self-censorship, better discriminate between groupthink and other group processes. It is difficult to accept the argument that the presence of characteristics common to most problem-solving groups is supportive of the existence of groupthink.

Janis did not accept this weak form of groupthink. To the contrary, he wrote the following in 1989:

> It does not suffice merely to see if a few of the eight telltale symptoms of groupthink can be detected. Rather, it is necessary to see if practically all the symptoms were manifested and also to see if the antecedent

conditions and the expected immediate consequences—the symptoms of defective decision making—are also present.

Recent research has continued to show little empirical support for the groupthink phenomenon. For example, in 2000, Park conducted a comprehensive empirical investigation of Janis's model, including all 24 variables (antecedent conditions, symptoms of groupthink, signs of defective decision making, and outcomes). Of 23 predictions drawn from the groupthink model, only 2 were confirmed. Conversely, 7 of the 23 relationships were significant opposite the direction predicted.

As another example, in 1999, Choi and Kim examined groupthink and team activities in 30 organizational teams faced with impending crises. They found that groupthink symptoms consisted of two factors. Contrary to groupthink predictions, one of those factors (termed "group identity") was significantly positively related to team performance, whereas the other factor (termed "concurrence seeking") showed an insignificant negative relationship to performance.

Addressing what is probably the most-cited example of groupthink, the space shuttle *Challenger* disaster, Maier (who developed a popular documentary on the topic) noted in 2002 that new evidence regarding the disaster and further analysis of past evidence demonstrates convincingly that the disaster "emphatically is not an example of groupthink." He noted that two of groupthink's defining features, the conviction of invulnerability and the illusion of unanimity, were conspicuously absent. Maier discussed evidence that the decision to launch was driven by uncertainties rather than perceived infallibility and that certain actions were taken only when it was clear that opinions would not be unanimous.

In short, research continues to find very little empirical support for the groupthink phenomenon. Coupled with several theoretical critiques, such evidence calls into question the validity of the phenomenon. Elsewhere, it has been noted that others have attached a variety of possible meanings to groupthink, as indicated in what follows.

• Groupthink is overreliance on concurrence seeking.

This meaning, which seems to have gained considerable popularity, would grant no value added to groupthink. That is, overemphasis on concurrence seeking was a widely recognized phenomenon decades before Janis coined the term "groupthink" (e.g., by Schanck in 1932), and the groupthink phenomenon simply adopted this element. If groupthink is to have any unique explanatory value, it must have additional meaning beyond this sole concept.

• Groupthink is a complete sequence of clusters of dysfunctional characteristics that result in flawed decision making.

As noted earlier, Janis accepted this view of groupthink, which is labeled a "strong view" in this article. Quite simply, there is no support for this interpretation of groupthink. No study conducted has ever supported—or come close to supporting—this strong view. Instead, researchers have adopted weak versions in which any partial support, regardless of the number of disconfirming findings and whether or not that partial support relates to elements unique to the groupthink model, is hailed as further evidence of groupthink's validity. Furthermore, support for the posited groupings of groupthink characteristics derives from anecdotes, casual observations, and intuitive appeal rather than from rigorous research. There has been no full factor analysis of groupthink variables. Incomplete factor analyses (in which exploratory factor analysis was applied to variables within sets rather than to all variables in the model) support a simpler model than that presented by Janis. Path analysis of the complete groupthink model has revealed only a scattering of the relationships proposed by Janis. In addition, in 1992, Tetlock and colleagues used a quantitative method, the group dynamics q-sort, for analyzing historical cases of group decision making to compare Janis's case interpretations with other historical interpretations of the same cases and to test the causal linkages predicted by the groupthink model. A test of the causal sequence of the groupthink model yielded only limited support. For example, group cohesiveness and provocative situational context proved to be only weakly related to concurrence seeking, with structural and procedural faults playing larger roles. In sum, there is no solid support for the clusterings of characteristics posited by the groupthink phenomenon or for their proposed causal sequence.

• Groupthink is an undesirable constellation of characteristics resulting from highly cohesive groups.

Again, there is no empirical support for this meaning. Cohesiveness, supposedly the critical trigger in the groupthink phenomenon, has simply not been found to play a consistent role. Conceptually, there are many reasons to question this view of groupthink. First, the

benefits of group cohesion, including enhanced communication among members, heightened member satisfaction, decreased member tension and anxiety, and higher levels of task accomplishment, have long been recognized. Furthermore, any general statement about the dangers of cohesiveness ignores critical contingency variables. As just one example, the stage of group development may be important because members of a fairly cohesive group in the mature stage of group development are likely to be secure enough in their roles to challenge one another. Given all of this, the findings relating to cohesiveness in the groupthink literature are unsurprising in that they demonstrate convincingly that cohesiveness does not regularly lead to negative outcomes. More than 25 years ago, Flowers stated that a revision of Janis's theory, one that would eliminate cohesiveness as a critical variable, may be justified. Any such elimination would, of course, largely eviscerate the groupthink phenomenon. Janis viewed cohesiveness and an accompanying concurrence-seeking tendency that interferes with critical thinking as the central features of groupthink.

- Groupthink is any set of group processes that are antecedent to poor decision outcomes.

This view is tautological, suggesting that groupthink is all of those bad things that precede poor outcomes. It is not surprising, then, when poor outcomes are found to follow these bad things. Furthermore, because many group processes and characteristics leading to poor outcomes had been identified long ago, this meaning, in which groupthink most clearly adopts the form of a metaphor for dysfunction, essentially grants groupthink no status—and Janis no contribution—beyond that of providing a memorable label. Nevertheless, much of the popularity of groupthink may be attributable to adoption of this perspective. That is, there is apparently a reluctance to abandon a compelling metaphor that encourages focus on flawed group processes.

4. BASES FOR GROUPTHINK'S APPEAL

In the current authors' view, groupthink has gained the status of received doctrine. Groupthink has a seductive intuitive appeal, and vivid examples that are superficially consistent with the phenomenon can readily be recounted. But how did groupthink come to be so widely and enthusiastically accepted in the absence of compelling evidence? This section, based in part on Aldag and Fuller's earlier review in *Psychological Bulletin,* explores the bases for groupthink's appeal.

4.1. Implicit Assumptions

Groupthink's appeal appears to rest in part on a number of implicit assumptions. First, although groupthink assumes that groups are used primarily to enhance the quality of decisions, group decision making is often used specifically to increase the acceptance of a decision. Also, organizational members may choose to use groups for other reasons, including socioemotional benefits and accomplishment of members' own secondary goals (e.g., "hidden agendas").

Second, groupthink implicitly assumes that groups rationally pursue a unified goal. However, group processes are characterized by interplays of covert and overt motivations, concern not just about the current problem but also about residues of past problems and the anticipation of future problem-related interactions, and many other subtleties outside the typical focus of the rational model. Processes and outcomes that appear irrational to "objective" observers may in fact be functional from the point of view of individual members or even of the entire group. For instance, such processes and outcomes may serve to maintain the motivation of the group leader, help ensure the continued use of group processes, prevent the defection of certain group members, serve as a means for members to show allegiance to coalitions, and minimize the likelihood of the use of still other political acts.

A third implicit assumption is that actions taken to prevent or minimize groupthink will produce a more rational and effective group process. However, these remedies may have unintended consequences. That is, attempts to create a microrational decision process in an inherently political environment may prove to be misguided; other literatures have suggested the dangers of presenting an oasis of change in a nonreinforcing desert. The groupthink remedies are decision specific, whereas the causes for supposed group dysfunctions may be systemic. That is, the application of specific techniques to prevent or minimize group difficulties may amount to treating the symptom of a deeper problem.

Fourth, fundamental to groupthink is the assumption that a group's illusion of well-being is dysfunctional. However, in 1988, Taylor and Brown argued that such illusions actually promote many criteria of mental health, including the ability to care for others, be happy or

content, and engage in productive and creative work. They added that such positive illusions are especially useful when an individual receives negative feedback or is otherwise threatened and may be especially adaptive under these circumstances. Thus, this illusion of well-being may be most useful specifically in the hot decision situations described by Janis. Although Taylor and Brown focused on the individual, the processes that they detailed seem equally applicable to groups.

Such contentions are reinforced by consideration of the related literature of the self-fulfilling prophecy, referred to in some settings as the Pygmalion effect. The Pygmalion and similar expectation effects appear to be common in organizational settings. Taken as a whole, evidence concerning these related effects suggests that nonveridical perceptions may often yield positive consequences. The literature on optimism–pessimism leads to similar conclusions. That is, although pessimism is often found to be associated with more realistic worldviews, optimism typically yields superior outcomes. These arguments are not meant to imply that overly positive perceptions are necessarily desirable; they are only meant to imply that the opposite should not be automatically assumed.

A final groupthink assumption concerns causes of characteristic groupthink attitudes and behaviors. Rather than, or in addition to, stemming from individual psychological factors, conformity behaviors may be politically motivated. For instance, they may reflect compliance rather than internalization. Thus, many groupthink characteristics could be due to conscious actions by members to subscribe to political norms.

What the preceding implicit assumptions have in common is a beguiling, albeit unrealistic, simplicity.

The reliance on anecdotes in groupthink scenarios and case studies contributes to the availability heuristic. It does this in at least two ways. First, groupthink examples tend to be vivid and "catchy" and, thus, are salient and readily available in individuals' memories. Second, such examples play to individuals' preference for case, as opposed to base, data. A concrete instance of the appearance of groupthink symptoms in a fiasco may be seen as compelling evidence, especially in the absence of base data.

4.2. Implicit Theories of Group

The groupthink phenomenon is consistent with implicit theories of groups. The roles of implicit theories in influencing responses have been widely documented. One important characteristic of implicit theories is their "gap-filling" function. In the context of groupthink, individuals observing a situation in which some groupthink characteristics are present may assume the existence of others. Furthermore, research shows that feedback given to group members or observers about group performance affects the characteristics ascribed to those groups. For example, feedback about group process has been shown to affect evaluations of outcomes, and feedback about outcomes has been shown to affect evaluations of process. Individuals who are told that a group has performed poorly are more likely to report instances of "poor" interaction processes such as lack of willingness to hear other members' views. Thus, the focus on poor decision outcomes in groupthink research may lead to reports of poor group functioning.

4.3. Illusory Correlation

A focus on only the conjunction of groupthink characteristics and negative outcomes invites illusory correlation. In this sense, the groupthink phenomenon is similar to the "Friday the 13th" phenomenon; only the yes–yes (i.e., groupthink/Friday the 13th–poor outcomes/bad luck) cell is considered. If the yes–yes cell is not empty, support for the phenomenon is inferred. In fact, of course, support for the phenomenon requires examination of all cells. Furthermore, when research is used to explain a poor decision, the process used to "validate" groupthink nearly guarantees support, however partial. That is, in the aftermath of a fiasco, transcripts and other documents, videos, and casual observation may be sifted for "evidence" of groupthink. If that search unearths comments that deal with conformity, caution, or closed processes, a "hit" is registered in the yes–yes cell and the presence of groupthink is announced. However, some reference to such terms is nearly inevitable in the mass of evidence surrounding a fiasco.

4.4. Framing

The negative language of groupthink (e.g., "victims of groupthink," "defects of groupthink") and the focus on error can invite distortions in responses caused by scale use tendencies and related psychometric difficulties and can also result in framing effects. Individuals (whether research participants or researchers) presented with negatively framed terminology may adopt the readily available negative frame and respond accordingly. Therefore, even simple attempts by the participants to give responses that are consistent with the tone of the

questions would result in negatively oriented responses. On a more fundamental level, this framing has resulted in a focus on error rather than on decision quality per se. Janis noted in 1982 that he began studying fiascoes "for the purposes of studying sources of error in foreign policy decision making." However, just as there is more to the performance of a football team than the absence of fumbles and interceptions, there is more to group decision quality than the absence of error. A focus on negative outcomes of group processes may divert attention from group synergies. One example is the assembly effect bonus, which is productivity that exceeds the potential of the most capable member as well as the sum of the efforts of group members working separately. Thus, one may learn little about superior group performance by focusing solely on fiascoes. Instead, a focus on decisions with a broad range of outcomes, including superior performance, is necessary.

In the current authors' view, groupthink owes much of its popularity to a characteristic it shares with other popular concepts such as total quality management. A core concept with some validity (e.g., statistical quality control) is incorporated as an element of a broader renamed concept (e.g., total quality management). Support for the core concept is subsequently treated as confirmation of the broader concept and, by association, for its various elements. In the case of groupthink, a core concept with some validity (i.e., overemphasis on concurrence seeking may be harmful) is subsumed in a complex, essentially deterministic model. Subsequent instances of that core concept are then presented as evidence for the validity of the broader groupthink phenomenon.

5. IMPLICATIONS AND SUGGESTIONS FOR FUTURE RESEARCH AND THEORY

Groupthink has stimulated research on group dysfunctions, provided links to other literatures such as stress and vigilance, emphasized potentially important variables in group decision making, and encouraged policymakers to take remedies for excessive concurrence seeking seriously. Indeed, Janis's recommendations for "remedies" for groupthink are themselves an excellent compilation of approaches to help preclude group dysfunction.

However, group problem solving is a complex and challenging phenomenon. Although the groupthink model has played an important role in generating interest in group problem-solving processes and in acting as a catalyst for associated theory and research, it has not incorporated three decades of theory and research, has received limited empirical support, and is restrictive in scope. Recent theory and research, as well as critical evaluation of the model, suggest that more comprehensive models are necessary to guide researchers and practitioners in dealing with the group decision phenomenon.

It is not unusual for theories to generate initial widespread interest and enthusiasm and then to meet with subsequent revision, rejection, or reaffirmation. However, groupthink has generally resisted dispassionate reevaluation. This is perhaps due in large part to the raw intuitive appeal of the phenomenon as well as to the fact that studies of groupthink have often been searches for confirmation. Rigorous evaluation of the phenomenon is further rendered difficult by the fact that a variety of views of groupthink have emerged, as have contrasting positions on what level of evidence is needed to indicate support.

Similarly, popular acceptance of groupthink has been extraordinary. Perhaps this is understandable: Groupthink has served as a vivid bogeyman that can be readily summoned to illustrate the dangers of overemphasis on concurrence seeking, and it continues to serve its purpose. This article has provided a variety of potential reasons for this enthusiastic acceptance and has shown that they are consistent in yielding illusory support.

Although attention to groupthink has yielded some clear benefits over the past three decades or so, there is growing consensus that the phenomenon lacks conceptual and empirical support, is restrictive in scope, and is inappropriately deterministic. As such, future research should seek to build on decades of groupthink and other research and theory on group processes and outcomes to yield more encompassing, less causally constrained perspectives. New theories and models should generalize to a broad range of group types and situations (e.g., including self-managing groups and virtual groups), incorporate additional relevant variables (e.g., leader motivations, stage of group development, power relationships), and recognize that groups, as well as group leaders, may have multiple goals (e.g., sustained motivation of the group leader, continued use of group processes, prevention of defection of group members). Finally, the research should avoid selection on the dependent variable; instead, it should consider instances of both favorable and poor group

outcomes and should seek to be theory building rather than theory confirming. In this way, the energy, enthusiasm, and intellectual and other resources that have been directed toward the groupthink phenomenon for more than three decades may yield broader, more inclusive, and more valid models and theories of group decision processes.

6. PRESCRIPTIONS FOR EFFECTIVE GROUP PROBLEM SOLVING

As this analysis has indicated, the current authors believe that the topic of group problem solving should be viewed more comprehensively than within the confines of the restrictive groupthink phenomenon. Taking into account the broader literature on group problem solving, it is clear that there are legitimate threats to group processes beyond—and sometimes inconsistent with—those suggested by the groupthink model. With this in mind, the authors offer the following advice to increase the effectiveness of group problem solving.

The first recommendation is to pay careful attention to group process and provide structure and guidance to ensure that the full process proceeds effectively. There are five stages to the problem-solving process: problem identification, generation of alternatives, evaluation of alternatives, decision making, and decision implementation and control. It is important that a group proceed systematically through these stages, allotting appropriate time and attention to each stage. In addition to proper structuring of the process by the group leader, some tools (e.g., the nominal group technique) and computerized decision aids (e.g., decision insight systems) may also prove to be valuable for this purpose.

A second recommendation is to consider the level of conflict in the group. The leader must be able to assess the degree of conflict and strive to bring the group toward an optimal level, where different viewpoints are encouraged and accepted but where the differences are viewed as adding information and contributing to the discussion. In any group, or at any time, conflict might have to be reduced, stimulated, or maintained. If conflict levels are very low, productive dissent might be needed. That is, as suggested by Schanck and others, group decisions may suffer if there is a strong pressure to conform and remain silent against the dominant view. In such cases, techniques suggested by Janis (e.g., the devil's advocate, "second chance meetings") might be useful. In addition, leaders can form subgroups that take different sides of an issue and can encourage norms that all employees, regardless of their positions, should fully express their concerns, doubts, and ideas, emphasizing that it is essential to examine issues critically and to find flaws and problems before competitors or customers do. Another approach to encourage productive dissent is to form task forces, made up of employees with various perspectives, outlooks, organizational positions, and backgrounds, to examine major problems and offer recommendations. Conversely, too much conflict will also result in poor group process and decisions. In such cases, it may be useful to encourage the group to focus on larger goals, directly intervene to alleviate conflict, separate conflicting parties, work to improve communication among members, train members in finding integrative solutions, or bring in third-party mediators.

Third, the discussion suggests that it is important to consider many outcomes of group problem solving rather than focusing just on short-term decision quality; decision implementation and commitment to decisions may be as important as the decision itself. Future motivation of the group and leader, future use of the group, and member affective responses may be important outcomes for the group and/or its members. The leader must recognize trade-offs among multiple goals, sometimes subordinating short-term decision quality to goals such as maintaining member satisfaction and enhancing the likelihood of future use of the group.

Fourth, caution must be exercised in the use of any individual technique. In an organization, a group functions within a broader environment, that is, the culture of the organization. As such, a specific remedy applied to a group may be ineffective if it contradicts the dominant cultural norms in the organization. Therefore, attempts to improve group functioning must be viewed within the broader context and should explicitly consider systemic issues. It may be more appropriate to assess and perhaps alter organizational culture or contingencies of reinforcement than to implement specific techniques in the hopes of overcoming the consequences of the group's decision environment.

Finally, it is important for the group leader to recognize and manage the complexity and challenges of group problem solving. The leader must strive to keep an open mind, avoiding limiting assumptions and models—no matter how alluring—and maintaining broad behavioral repertoires to permit flexible responses to the rich contingencies influencing group processes and outcomes.

See Also the Following Articles

Conflict within Organizations ■ Groups, Productivity within ■ Intergroup Relations and Culture ■ Interpersonal Conflict

Further Reading

Aldag, R. J., & Fuller, S. R. (1993). Beyond fiasco: A reappraisal of the groupthink phenomenon and a new model of group decision processes. *Psychological Bulletin*, *113*, 533–552.

Callaway, M. R., & Esser, J. K. (1984). Groupthink: Effects of cohesiveness and problem-solving procedures on group decision-making. *Social Behavior and Personality*, *12*, 157–164.

Choi, J. N., & Kim, M. U. (1999). The organizational application of groupthink and its limitations in organizations. *Journal of Applied Psychology*, *84*, 297–306.

Flowers, M. L. (1977). A laboratory test of some implications of Janis's groupthink hypothesis. *Journal of Personality and Social Psychology*, *33*, 888–895.

Fuller, S. R., & Aldag, R. J. (1998). Organizational Tonypandy: Lessons from a quarter century of the groupthink phenomenon. *Organizational Behavior and Human Decision Processes*, *73*, 163–184.

Janis, I. L. (1971, November). Groupthink. *Psychology Today,* pp. 43–46, 74–76.

Janis, I. L. (1972). *Victims of groupthink*. Boston: Houghton Mifflin.

Janis, I. L. (1982). *Groupthink* (2nd ed.). Boston: Houghton Mifflin.

Janis, I. L. (1989). *Crucial decisions: Leadership in policymaking and crisis management*. New York: Free Press.

Janis, I. L., & Mann, L. (1977). *Decision making: A psychological analysis of conflict, choice, and commitment*. New York: Free Press.

Longley, J., & Pruitt, D. G. (1980). Groupthink: A critique of Janis's theory. In L. Wheeler (Ed.), *Review of personality and social psychology* (pp. 507–513). Beverly Hills, CA: Sage.

Maier, M. (2002). Ten years after *A Major Malfunction . . .*: Reflections on "The *Challenger* syndrome." *Journal of Management Inquiry*, *11*, 282–292.

Park, W. (1990). A review of research on groupthink. *Journal of Behavioral Decision Making*, *3*, 229–245.

Park, W. -W. (2000). A comprehensive empirical investigation of the relationships among variables of the groupthink model. *Journal of Organizational Behavior*, *21*, 873–887.

Paulus, P. B. (1998). Developing consensus about groupthink after all these years. *Organizational Behavior and Human Decision Processes*, *73*, 362–374.

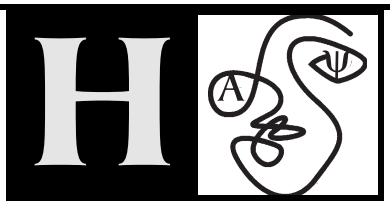
Schanck, R. L. (1932). A study of a community and its groups and institutions conceived as behaviors of individuals. *Psychological Monographs, 43*(2, whole No. 195).

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social–psychological perspective on mental health. *Psychological Bulletin*, *103*, 193–210.

Tetlock, P. E., Peterson, R. S., McGuire, C., Chang, S., & Feld, P. (1992). Assessing political group dynamics: A test of the groupthink model. *Journal of Personality and Social Psychology*, *63*, 403–425.

't Hart, P., Stern, E. K., & Sundelius, B. (1997). *Beyond groupthink: Political group dynamics and foreign policymaking*. Ann Arbor: University of Michigan Press.

Whyte, G. (1989). Groupthink reconsidered. *Academy of Management Review*, *14*, 40–56.



Health and Culture

George D. Bishop
National University of Singapore, Singapore

1. The Concept of Culture
2. Culture and Health-Related Cognition
3. Culture and Health-Related Behavior
4. Culture and Physiological Responses
5. Integrating Culture into Health Psychology
6. Conclusion
Further Reading

GLOSSARY

balance model A model that describes health in terms of a harmonious balance between different elements or forces in the body.

culture Guidelines that are inherited from the social environment that guide how people in a particular cultural group view the world, respond to it emotionally, and behave in it.

machine model A model of the body as being an internal-combustion or battery-driven machine.

"plumbing" model A model that views the body as being a series of cavities or changes connected to each other through pipes or tubes.

This article examines the interrelationship of culture and health with emphasis on the ways in which culture influences health outcomes through its effects on health-related thoughts and behaviors as well as physiological responses. Examples are given of these influences and the article concludes with a discussion of how cultural influences can best be integrated into research and theory in health psychology.

Considerations of culture are becoming increasingly prevalent in psychology, and health psychology is no exception. This article explores the relationship between culture and health and ways in which considerations of culture are important in the development of a theoretically and empirically sound health psychology. Specifically, culture plays a key role in how people understand states of health and illness as well as the types of health-related behavior in which they engage. In addition, there is growing evidence that culture influences physiological responses. As such, considerations of culture are essential to understanding various aspects of health.

1. THE CONCEPT OF CULTURE

Many definitions have been given for culture. Most, however, view culture as a set of guidelines that people inherit from their social environments that guide how they view the world, respond to it emotionally, and behave in it relative to the natural environment, other people, and supernatural forces or gods. As such, culture provides a set of lenses that shape experience. These "lenses" are transmitted from one generation to the next in various ways and provide the basis for shared experience within the cultural group. One of the ways in which culture shapes experience is through the provision of categories for dividing and organizing experience. In the case of health, such categories are key to understanding how people experience health

©2004 Elsevier Inc.
All rights reserved.

and illness and how they report that experience to others, such as health professionals.

It is important to note that culture is not something that is uniform within a society; rather, any society is certain to have multiple cultures within it. These cultures may be differentiated on the basis of ethnicity, social class, gender, occupation, and any number of other attributes that are associated with shared experience within a group that differentiates it from other groups. Furthermore, cultures change over time as different cultural groups interact, conditions change, technology develops, and so on. Hence, culture is best seen as dynamic rather than static.

2. CULTURE AND HEALTH-RELATED COGNITION

One of the key ways in which culture influences health is through health-related cognition.

2.1. Variation

Variation in health and illness concepts between cultures is obvious even to the most casual observer. These variations are exhibited in a wide variety of ways, but four of these are particularly striking.

2.1.1. The Basic Nature of Health and Illness

First, there is substantial variation between cultures in conceptions of the basic nature of health and illness. For example, in Western societies placing a strong emphasis on the scientific understanding of health and illness, these concepts are understood in natural terms with disease viewed as pathology due to natural causes such as germs, genetic aberrations, or injury. In some other societies, however, as well as among some minorities in Western societies, illness may be understood more in interpersonal or supernatural terms. For example, illnesses may be seen as being the result of violations of social norms or standards of personal behavior, particular emotions (e.g., anger or jealousy), or transgressions of moral or religious taboos.

2.1.2. Beliefs about Anatomy and Physiology

Second, beliefs about anatomy and physiology often show marked cultural variation, with cultures often differing in their basic model of the body and what constitutes health and illness. For example, a number of forms of traditional medicine, including Chinese traditional medicine as well as traditional medical beliefs in India, Latin America, and the Caribbean, hold to the balance model. This model describes healthy functioning of the body in terms of a harmonious balance between different elements or forces within the body. Illness is believed to be caused by imbalance, which must be rectified for the person to return to health. In contrast, in Western societies the body is often conceptualized in terms of the "plumbing" model and/or machine model. The former views the body as being a series of cavities or chambers (e.g., stomach and chest) connected to one another by pipes or tubes. In this model, health is maintained by an uninterrupted flow of various substances (e.g., blood, air, food, and urine) between cavities or between cavities and the exterior of the body. By the same token, disease is the result of blockage. The machine model, by contrast, views the body as a machine that operates according to the principles of physics and chemistry and that may run down or break down, requiring a doctor to repair it.

2.1.3. Concepts and Terminology

Third, cultures often differ markedly in the concepts and terminology used to describe various conditions. For example, symptoms often diagnosed as depression in the West may be interpreted as neurasthenia (nervous exhaustion) in China. Also, conditions defined in one culture may be undefined in another. For example, traditional Mexican American medical beliefs define conditions such as *susto* (fright), *empacho* (surfeit), *caida de mollero* (fallen fontenelle), and *mal de ojo* (evil eye), which have no direct referents among conditions known to medical professionals or Anglo-Americans.

2.1.4. Underlying Structure of Illness Representations

Finally, the underlying structure of illness representations often differs between cultures. Research has shown that the mental representations that people have of different diseases show a characteristic organization in memory that differs from culture to culture. For example, research has consistently found that disease representations among Anglo-Americans are organized along dimensions of contagiousness and severity. In contrast,

disease representations among Mexican Americans appear to be organized according to concepts of "hot" and "cold" and whether a disease is associated with children or old people, and those of Chinese Singaporeans are organized according to a combination of dimensions deriving from Chinese and Western medicine.

2.2. Implications

These cultural differences in health-related cognition have a number of important implications. First, the cognitive representations that people have of diseases may well have a strong influence on their interpretation of symptoms and bodily symptoms. The existence of specific disease concepts makes it possible and perhaps even likely that bodily experience will be interpreted in line with that concept, whereas the absence of that concept would foreclose that possibility. For example, the fact that Mexican Americans have available to them the concept of empacho makes it possible for them to experience this condition, whereas members of other groups not possessing this concept would not be able to experience it. Similarly, the prominence of the concept of neurasthenia in China and of depression in North America makes it more likely that the similar symptoms associated with these two conditions would be interpreted, and quite likely experienced, differently.

In addition, health-related cognitions have implications for what people do when experiencing distress. Illness representations provide a definition for when bodily distress represents illness and requires assistance from a medical professional or other healer. It has been noted, for example, that one of the reasons why heart attack victims delay seeking treatment is that they misinterpret the symptoms and underestimates the seriousness of the situation. Furthermore, once the person has determined that help is needed, mental representations of illness direct the person's help-seeking behavior. Research with Chinese Singaporeans has shown that whether a person seeks help from a practitioner of Western scientific medicine or Chinese traditional medicine depends on the extent to which the person endorses illness beliefs associated with Chinese traditional medicine. Finally, once help is sought, people's mental representations of their condition are likely to strongly influence whether they stay in treatment. Research with hypertensives, for example, has shown that the cognitive "model" they have of hypertension and its relationship to specific symptoms is a key determinant of whether they stay in treatment.

3. CULTURE AND HEALTH-RELATED BEHAVIOR

A second important way in which culture influences health is through behavior.

3.1. Behavior and Disease Risk

The relationship between behavior and health is well established and there is strong evidence for the importance of behavioral risk factors such as smoking, diet, exercise, and alcohol use. It is also obvious to even the most casual observer that these behaviors often differ dramatically between cultural groups. For example, alcohol use is very prevalent in some countries, such as Russia, Ireland, France, and Italy, but much lower in other countries, such as predominantly Muslim countries in which alcohol use faces religious prohibition. Also, dietary habits differ substantially between cultures with regard to what is defined as food as well as which foods are allowed or prohibited, how foodstuffs are prepared, and overall patterns of eating. For example, although dog meat is eaten in some societies, most Americans and many others are appalled at the idea of considering dogs to be a source of food. Also, religion often dictates what foods are allowed. Orthodox Hindus, for example, are forbidden to eat any kind of animal, whereas Jews and Muslims are prohibited from eating pork, and any meat they do eat must be slaughtered in specific ways.

Cultural differences in behavior have an important impact on diseases prevalent in different countries and among different cultural groups. For example, research on cultural differences in diet has shown that the "Western" diet, with its high saturated fat content, is associated with high levels of coronary heart disease (CHD), whereas the low-fat Japanese diet is associated with low levels of CHD. Furthermore, the "Mediterranean" diet, which is high in monounsaturated fats and antioxidants, has been found to be protective against CHD. Similarly, elements of the Japanese diet have been associated with the development of stomach cancer and the habit of chewing a mixture of tobacco, betel nuts, and other substances in countries such as India and Afghanistan has been associated with high levels of oral cancer. Also, studies of alcohol consumption have shown that cross-cultural differences in alcohol consumption are strongly related to the incidence of diseases such as liver disease, CHD, stroke, hypertension, and diabetes. Finally, the

prevalence of HIV infection and AIDS as well as other sexually transmitted diseases is strongly related to cultural attitudes toward sexuality and the acceptance of promiscuity, commercial sex work, and alternative sexual practices.

3.2. Help-Seeking

In addition to culture's role in the etiology of disease, it is also important to note cultural differences in help-seeking and the role of culture in treatment. As noted earlier, cultural differences in cognitive representations of disease are likely to influence whether a person seeks treatment for specific symptoms. In addition, there are also cultural differences in reasons for seeking treatment and the kind of treatment that is considered to be appropriate. For example, it has been noted that for Italian Americans, seeking medical care tends to be triggered by an interpersonal crisis or when the person perceives that the symptoms interfere with personal relationships. In contrast, Irish and Anglo-Americans tend to seek help when symptoms are perceived to interfere with work or physical functioning. Furthermore, when seeking help different cultural groups may present their symptoms differently and adopt a different "language of distress." Research has found that Irish Americans tend to underplay their symptoms, taking a stoical attitude toward them, whereas Italian Americans tend to present their symptoms in a more voluble, emotional, and dramatic manner. These differences, in turn, may have an influence on diagnosis and the treatment received.

After the person has sought treatment, culture may play an important role in the interaction with the health care practitioner. Due to migration, most societies are becoming increasingly multicultural, with the result that it is commonly the case that patients and health care practitioners may come from different cultural backgrounds. When this is the case, differing cultural assumptions may prove a barrier in communication as well as in the acceptance of medical advice. For example, it has been noted that Native Americans tend to underuse medical services that are available to them and often do not follow medical recommendations even when they do utilize services. Among the reasons for this are conflicts between scientific and traditional Native American understandings of disease as well as the history of oppression of Native Americans and differences in communication styles between nonnative health care practitioners and Native clients. The end result may be distrust of the nonnative health care practitioner unless the practitioner makes specific efforts to understand Native beliefs and also approaches Native clients in a manner that employs both verbal and nonverbal behavior that is culturally appropriate.

4. CULTURE AND PHYSIOLOGICAL RESPONSES

In addition to the often very evident effects of culture on beliefs and behavior, there is also a growing body of evidence suggesting cultural differences in physiological responses that may have important health implications. There is often a tendency to assume that physiological functioning and disease processes are the same across cultural groups. However, this may not always be the case. For example, numerous studies have shown significant relationships between personality traits such as trait anger and hostility and cardiovascular responses to certain stressors. However, there is also evidence that this relationship may vary according to ethnicity. In one study, correlations between hostility and cardiovascular responses to stress that were negative among Caucasian participants were positive among African American participants and vice versa. Also, there is evidence that Asian Indians show a different pattern of cardiovascular responses to stress than do Chinese or Malays. Other research has shown that the relationship of socioeconomic variables and hostility to cardiovascular reactivity to stress and left ventricular mass is different between Caucasians and African Americans.

The reasons for these differences are not clear and they may be the result of either cultural differences between groups or genetic differences. There is evidence, however, that even genetic influences may vary by cultural group. For example, studies of the serotonin transporter gene have shown that although polymorphisms of this gene are related to both personality and cardiovascular responses to stress, this relationship differs between African Americans and Caucasians. This suggests an interaction between culture and genetics in physiological responses to stress that may have important implications for understanding health differences between groups.

5. INTEGRATING CULTURE INTO HEALTH PSYCHOLOGY

It is clear from the previous discussion that culture and health are closely intertwined, with cultural factors

having a strong influence on health through cognitive, behavioral, and physiological pathways. However, the study of cultural factors in health psychology has only recently begun, and there is a great deal more that needs to be done to fully integrate cultural considerations into health psychology.

Research on cultural differences has most commonly involved making comparisons between specified cultural groups, an approach sometimes referred to as the "packaged" approach. The assumption is that differences between cultural groups are the result of differing cultural characteristics. However, this approach suffers from the fact that it ignores potential genetic and other physical differences between groups and does not reveal which aspects of culture are responsible for the differences observed.

Because of the limitations of the packaged approach, cross-cultural psychologists strongly advocate going beyond the superficial aspects of existing cultural groupings by examining the underlying dimensions of culture. Any number of dimensions might be identified, but attention has focused primarily on dimensions identified in classic research by Hofstedte (individualism–collectivism, masculinity–femininity, certainty avoidance, and power distance) as well as those based on cultural values. Only a few studies have examined the relationship between dimensions of culture and health-related variables, but results have been promising. For example, studies of Chinese values and their relationship to health have shown that the values of social integration and reputation are significantly related to both alcohol and tobacco consumption as well as to diseases such as cerebrovascular disease, heart disease, and cancer.

6. CONCLUSION

Culture plays a significant role in health through several different mechanisms. Cultures differ in their basic understanding of the phenomena of health and illness as well as in their concepts of specific health conditions. These differences have important implications for the interpretation of symptoms and bodily symptoms and also for the kinds of help that people seek when experiencing distress and whether they stay in treatment once they seek help. Culture also plays a significant role in health-related behavior. Substantial cultural differences exist with regard to behaviors such as dietary practices, smoking, alcohol use, sexual practices, and exercise, which are known to have a significant impact on risk for various diseases. Also, cultural groups differ in their propensity to seek medical attention and cultural factors figure prominently in the type of care that people seek and in their relationship to health care providers. Finally, there is increasing evidence of cultural differences in physiological mechanisms related to disease. Although the evidence for the importance of culture in health is strong, there is a very real need to go beyond the study of differences between cultural groups to the examination of how underlying dimensions of culture are predictive of health-related cognition, behavior, and physiological responses.

See Also the Following Articles

Health Promotion in Schools ■ Health Psychology, Cross-Cultural ■ International Classification of Diseases (WHO) ■ Women's Health

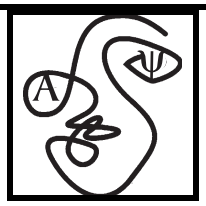
Further Reading

Angel, R., & Thoits, P. (1987). The impact of culture on the cognitive structure of illness. *Culture, Medicine, and Psychiatry*, *11*, 465–494.

Bishop, G. D. (1991). Understanding the understanding of illness: Lay disease representations. In J. A. Skelton, & R. T. Croyle (Eds.), *Mental representations in health and illness*. New York: Springer-Verlag.

Helman, C. G. (1995). *Culture, health and illness: An introduction for health professionals* (3rd ed.). Oxford, UK: Butterworth Heinemann.

Kazarian, S. S., & Evans, D. R. (Eds.). (2001). *Handbook of cultural health psychology*. San Diego: Academic Press.



Health Promotion in Schools

Ellis P. Copeland and Franci Crepeau-Hobson

University of Northern Colorado, Greeley, Colorado, USA

1. Introduction
2. Child and Adolescent Health Concerns
3. Health Promotion
4. Coordinated School Health Programs
5. Summary
   Further Reading

GLOSSARY

health A state of complete physical, mental, spiritual, and social well-being, not merely the absence of disease or infirmity.

health-promoting school A place where all members of the school community work collaboratively to provide students with integrated and positive experiences and structures that promote their health.

health promotion A combination of behavioral, educational, social, spiritual, economic, and environmental efforts that support the development, maintenance, and enhancement of behaviors and lifestyles conducive to emotional and physical well-being.

indicated prevention programming Interventions designed for children or families demonstrating danger signs (e.g., multiple risk factors, substance abuse, and neglect)and that are more intensive and intrusive than other prevention programs.

initiative The dedication of cumulative effort over time to achieve any goal composed of the elements of intrinsic motivation, concentration, and challenge.

mental health The successful performance of mental function, resulting in generative activities, fulfilling relationships, and the ability to cope with adversity.

selective prevention programming Interventions targeted at high-risk populations. The target can be children, families, or a high-risk school. The focus is most often on reducing risk factors or targeted behavior changes.

social emotional learning Knowledge, habits, skills, and values that provide a foundation for a child's academic, personal, social, and civic development and are necessary for success in school and life.

universal prevention programming Interventions designed for the entire population, such as a districtwide program designed to prevent substance abuse and/or violence.

The most serious and expensive health and social problems in the United States today are caused by negative behavioral patterns. Unhealthy behavioral patterns have adversely affected the physical, emotional, and mental health of youth, resulting in problems with alcohol, drugs, diet, anger management, and poor impulse control. These health problems and unhealthy behaviors are largely preventable and studies of health education and promotion in schools have demonstrated that health education programming has effectively reduced the prevalence of health risk behaviors among youth. Individual and family interventions have shown some success, but programs that champion a seamless, comprehensive (kindergarten–12) system of health promotion demonstrate great promise.

©2004 Elsevier Inc.
All rights reserved.

1. INTRODUCTION

Despite advances in medical sciences, today's youth face complex health challenges that are especially daunting. A significant number of youth enter school with "pre-existing" health conditions (e.g., asthma and diabetes) that, without appropriate intervention, can affect both academic and social progress. By adding communicable diseases, malnutrition, behavior-related conditions (e.g., exposure to smoking, eating disorders, substance abuse, and sexually transmitted diseases including HIV), violence, suicide, and accidents and injuries to the mix, one quickly realizes that there is a formidable problem. In the past, schools and other health and social services played a major role in impacting the health of young people through their involvement in immunization, screening, and educational programming. However, schools' involvement in health promotion has been sporadic and not as coordinated as previous immunization efforts.

Added to the health factors are the multiple environmental risk factors that are often antecedents of a variety of disorders. An increase in health and behavior problems can be linked to significant family, school, and neighborhood changes that have occurred during the past few decades. These changes include a breakdown in the traditional family and neighborhood; a devaluing of parenting, teachers, and religious life; damaging television and media messages (e.g., sex, violence, marketing of unhealthy dietary behaviors, and promotion of immediate gratification); lack of positive role models and a reduction in meaningful contact with positive adult role models; and an increase in inadequate housing and unsafe environments (often inclusive of the school environment).

Problems with regard to living conditions and the environment have taken their toll and, in large part, led the Centers for Disease Control and Prevention (CDC) to use the Youth Risk Behavior Surveillance System (YRBSS) to monitor the following six categories of priority health risk behaviors among youth and young adults every 2 years since 1991:

- Behaviors that lead to intentional (e.g., violence) and unintentional injuries
- Smoking/tobacco use
- Alcohol consumption and other drug use
- Sexual behaviors leading to disease, HIV infection, or unintended pregnancy
- Poor nutrition/unhealthy dietary behaviors
- Lack of/inadequate physical activity

These are all behaviors that can increase during childhood. The YRBSS monitoring system enables states to compare their data to national data and trends.

Relatedly, vital statistics indicate that 70.6% of all deaths among youth and young adults aged 10–24 years result from motor vehicle accidents, unintentional injuries, homicide, and suicide. Analysis of this national school-based survey indicates that many high school students engage in behaviors that increase their risk for death from one of these four causes. For example, 14.1% of the youth surveyed had rarely or never worn a seat belt during the 30 days preceding the survey, and 8.8% had attempted suicide during the 12 months prior to the survey.

The leading causes of mortality and morbidity among all individuals in the United states are related to health risk behaviors that are frequently established during youth: tobacco use, alcohol and drug use, unhealthy dietary behaviors, physical inactivity, etc. Thus, a realistic analysis of the physical, behavioral, and mental health of youth appears to be in order.

2. CHILD AND ADOLESCENT HEALTH CONCERNS

2.1. Physical Health

In 2001, almost 14% of high school students surveyed were at risk for becoming overweight, whereas 10.5% already fit this category. The problem is greater for males (14% overweight) than females (7% overweight). Fewer than 22% consumed the recommended number of servings per day of fruits and vegetables and milk, and an increasing number of schools include fast-food choices on their lunch menus. Furthermore, eating habits that begin in childhood contribute to numerous chronic diseases in adulthood (e.g., heart disease, cancer, and osteoporosis).

In addition to the unhealthy trends in dietary habits of youth in the United States, there has been a decrease in opportunities for physical activity, with slightly more than half (51.7%) of high school students enrolled in physical education (PE) classes. Less than one-third of these students attended PE classes daily—an increase from 25.4% in 1995 but far below the 1991 level of 41.6% attending daily. The percentage of students reporting no vigorous or moderate physical activity in the past 7 days is 9.5%. Combining lack of physical activity with increasingly unhealthy eating patterns, in addition to 2 hours or more of television and/or video

and computer games per day, results in multiple risk factors. Not surprisingly, children who participate in regular physical activity have more favorable cardiovascular disease risk factors than less active children.

The vulnerable ages for tobacco use are 10–18 years, when most users start smoking, chewing, or dipping and become addicted to the substance. Very few individuals begin to use tobacco as adults; almost all first use occurs prior to graduation from high school. Slightly more than one in three persons use tobacco by age 18 (33.9% report using tobacco in the past 30 days). The earlier a young person begins to use tobacco, the more likely the person is to abuse the substance as an adult. A surgeon general's 1994 report directly linked the deaths of 434,000 million Americans each year to tobacco usage. Because both the duration and the amount of tobacco use are related to eventual chronic health problems, early interventions with youth concerning tobacco use must be stressed as a physical health issue.

2.2. Drugs and Alcohol

Underage drinking is a serious problem in the United States. Alcohol is the drug most commonly used and abused by children aged 12–17 years. Youth in the United States are reportedly drinking alcohol at an earlier age than ever before, typically beginning in early adolescence (approximately age 13). A recent survey found that more than half of eighth graders and 80% of high school seniors had used alcohol sometime in their lives, and that 25% of eighth graders and more than 60% of high school seniors had been drunk at some time. The percentage of middle and high school students who had five or more drinks of alcohol at one time (within a few hours) on one or more of the past 30 days reached a high in 1997 at 33.4% and was 29.9% in 2001.

Health-related risks of teenage alcohol use are significant. Children who begin using alcohol before the age of 15 are four times more likely to become an alcoholic as an adult than those who start drinking at age 21. Research indicates that more than 40% of teens that start drinking at age 14 or younger become dependent on alcohol as adults. The risk of death from car accidents and drowning increases with alcohol use. Drinking large amounts of alcohol can result in alcohol poisoning, which can lead to coma or even death. Adolescents who drink are much more likely to attempt suicide, engage in violence against others, and become victims of violent crime.

Young people who drink alcohol are also more likely to engage in other risky behaviors, including illicit drug use: Youth who drink are 7.5 times more likely to use some type of illicit drug. The most widely used illegal drug is marijuana, despite declining usage. In 2000, almost half of all 7th–12th graders reported having tried marijuana sometime during their lives, which was down from 53% in 1997. The number of these kids who regularly use marijuana declined 13% during the same time period (approximately 21% reported being regular users in 2000).

Unfortunately, the downward trends observed in marijuana use have not been noted for other types of drugs, especially so-called club drugs. For example, since 1999, teen ecstasy (MDMA) use has increased 71%. More than 8% of high school seniors reported having used ecstasy at least once. Other club drugs, including rohypnol, methamphetamine, and LSD, are also gaining popularity. These drugs have sometimes been added to the beverages of unsuspecting females in order to sedate and sexually assault them. In addition, club drugs can produce a range of negative effects, including hallucinations, paranoia, brain damage and amnesia, and, in some cases, death. Numbers of emergency room visits by young persons under the influence of club drugs, especially ecstasy, increased approximately 60% from 1999 to 2000.

Drug and alcohol use increases risks in other areas as well. Adolescents who use drugs and/or alcohol are much more likely to be sexually active than youths who do not drink or get high. Furthermore, substance-abusing teens are more likely to start having sex at an earlier age and to have multiple partners. These youth are less likely to use protection/contraception. Despite these well-documented trends, very few prevention, treatment, and counseling programs deal with the combined problem of substance abuse and risky sexual behavior.

2.3. Sexual Health

Research indicates that slightly more than half of females and nearly two-thirds of males have had intercourse by their 18th birthday. Reported rates of sexual activity have declined among male teens in recent years. However, this trend has not been observed for females. In fact, increases in sexual activity for female teens have been noted for each year of age during the past several decades. The greatest increases have been noted for 14- to 16-year-old females, with twice as many of these girls engaging in sexual behavior as there were 15 years ago. This is concerning for several reasons. First, the younger females are when they first have sex, the more likely they

are to have had unwanted or involuntary first sex. Furthermore, the younger the age of first sexual intercourse, the greater the risk of unwanted pregnancy and sexually transmitted diseases and infections.

It is estimated that 40% of teen sex is unprotected. Relatedly, 7 in 10 teen pregnancies occur in teens not using contraception, resulting in approximately 500,000 babies born to teen mothers each year in the United States. The majority of sexually active youth say they used some form of contraception the first time they had sex, and 57.9% of youth report they used a condom during their most recent sexual encounter.

Substantial morbidity and social problems also result from the unprotected sexual behaviors of youth. Every year, more than 3 million teens—approximately 25% of those who are sexually active—contract a sexually transmitted disease. Teens appear to be especially at risk for contracting chlamydia and gonorrhea, with reported rates for adolescent females being much higher than those for sexually active adults. These infections are the leading causes of pelvic inflammatory disease, and chlamydia is a common cause of infertility, ectopic pregnancy, and perinatal complications.

Human papillomavirus, which is associated with most cases of cervical dyspasia (a precancerous condition of the cervix), and hepatitis B virus are also prevalent among sexually active teens. Some studies suggest that up to 15% of sexually active adolescent females are infected with at least one of these conditions. Rates of teen HIV are also high. It is estimated that approximately 25% of new HIV cases each year occur in young people ages 13–21. Most of these new cases are sexually transmitted. Further complicating these issues is the fact that teens are less likely to have health insurance or a primary health care provider; thus, screening, diagnosis, and treatment of sex-related health problems are often delayed. These delays may lead to medical complications that would otherwise be avoidable.

2.4. Mental Health

Today, children are faced with more challenges than ever before, but it is clear that the current system of providing services is keeping the prevalence of mental disorders in check but doing little to reduce the number (incidence) of youth affected. Professionals have estimated that more than 20% of the 70.4 million children aged 17 or younger in the United States experience the signs and symptoms of a *Diagnostic and Statistical Manual of Mental Disorders* disorder during the course of a year. However, 11% experience significant impairment and 5% experience extreme functional impairment. The surgeon general's 1999 report estimated that 20.9% of youth met diagnosable criteria ($\leq$70) on the Child Global Assessment Scale. Anxiety disorders (13.0) followed by disruptive disorders (10.3) and mood disorders (6.2) were the most common impairments (the rate is greater than 20% due to comorbidity). There is a link between child and adult disorders, but a significant number of youth "grow out" of their disorders and a significant number of adults develop a mental disorder later in life.

By examining self-report data, an argument can be made that these national estimates are low. For example, a 2001 CDC report found that the number of middle and high school students that reported an "attempted suicide during the past 12 months" was 8.8%. The percentage of students who, during the past 12 months, stated that they felt so sad or hopeless almost every day for 2 weeks or more in a row that they stopped doing some usual activities was 28.3%. The percentage of youth who reported that they had carried a weapon (gun, knife, or club) in the past 30 days was 17.4%. These behaviors and self-reports are consistent with mood disorders and disruptive disorders and appear to indicate that national estimates of incidence and prevalence are low.

On average, only approximately one-fourth of children in need of mental health care receive the help they need. The use of psychotropic medication for children and youth has increased sharply, with more than $1 billion spent in 1998 for these medications, approximately two-thirds of which was spent on stimulants and antidepressants.

Schools are important settings in which children's mental disorders can be recognized and addressed; research suggests that schools may function as the de facto mental health system for children and adolescents. Of those receiving mental health care, 70–80% receive that care in the school setting. Furthermore, 83% of schools report providing case management for students with behavioral or social problems.

Approximately 60% of the 1500 school-based health centers in the United States have mental health professionals on staff. With support from primary care providers, nearly 80 of these clinics provide crisis intervention services. Zins and Wagner recognized the full spectrum of health services while providing numerous examples of mental health activities, such as helping with transitions from elementary to middle school and providing instruction in stress or time management.

Although there has been extraordinary progress in improving public health (including mental health)

through medical science, more is known about how to treat mental illness than about how to prevent mental illness and promote health. There are numerous research-based programs aimed at curbing at-risk behaviors, such as violence and drug use, but a rigorous applied psychology of how to promote positive youth development is lacking. However, the psychologist's role is obvious given the link between learning and health.

3. HEALTH PROMOTION

Today's changing world, in which political and market forces and monetary gain often dictate human behavior, is not the most favorable environment for children to develop moral and ethical values. Many children find at the end of their school years that they may be equipped with an academic qualification but not with the coping skills to face a competitive and often dangerous world. It is in this context that "health" in schools has come to occupy a place as important as learning the three R's. Zins and Wagner defined health promotion as a combination of behavioral, educational, social, spiritual, economic, and environmental efforts that support the establishment, maintenance, and enhancement of behaviors and lifestyles conducive to overall emotional and physical well-being. The definition is inclusive and systemic, and it directly encourages the mobilization of community resources and the development of environments that will lead to healthy behavioral patterns during the formative years of childhood and adolescence.

The health problems and unhealthy behaviors observed in young people are largely preventable and studies of health education in schools have demonstrated that programming has effectively reduced the prevalence of health risk behaviors among youth. Given these data, it appears that school health programs are becoming one of the most effective means available to prevent major mental and physical problems as well as significantly reduce the adverse social problems that affect children and youth. School-based clinics first began to appear in the 1980s. These early programs were problem focused (e.g., alcohol and drug use and teen pregnancy) and reported some success. For example, interventions in middle schools reported significant reductions in smoking, obesity, alcohol, and marijuana use.

Because the problem behaviors targeted by these individual programs often occur together, share risk and protective factors, and can be addressed by similar interventions, there is support for more comprehensive and coordinated models of risk prevention and promotion of protective behaviors and environments. These comprehensive initiatives typically target multiple outcomes; span a number of years; and coordinate community, family, and school-based efforts.

Many schools have implemented programs targeting one or more categories of problem behaviors (e.g., violence and drug use). Unfortunately, adoption of multiple programs to address social and health issues can result in a variety of implementation problems. Often, multiple programs are poorly coordinated and compete with each other as well as core academics for limited instructional time and resources. Furthermore, these programs are often of short duration and are rarely sustained from year to year. They frequently lack the environmental supports at home and at school that allow for the maintenance of newly learned knowledge and skills.

Health promotion requires psychologists to be a part of a team of professionals from many fields working cooperatively with teachers, students, community members, religious leaders, and parents to provide comprehensive school health programs. These integrated efforts include such disciplines as health education, nursing, physical education, nutrition (food) services, counseling, and school psychology. The integrated approach better mobilizes community resources to enhance health-promotion efforts in the schools and society. Also, the comprehensive effort reaches virtually all children, creating an environment that builds on health in a more integrated manner.

Psychologists and other mental health professionals interested in introducing and gaining support for implementing a health-promotion program in their schools would be best served by beginning with a formal needs assessment. If such an assessment is not feasible, facilitating open discussion and active listening among staff is another strategy that can be utilized to delineate a hierarchy of concerns. After the needs have been identified, a system analysis is necessary to obtain an understanding of such things as the school board's philosophy, goals, and budget. Next, stakeholders, including school personnel, who are invested in health education should be identified. In 1995, Kelly suggested that internal support should be gathered before moving to external audiences. After input and approval of all involved school staff have been obtained, direct communication with external audiences can be initiated. Communication with community agencies, parents, churches and synagogues, hospitals, alternative health providers, and the business community is necessary to mange issues that may arise, including political groups that may perceive

health promotion as "too radical." Furthermore, direct communication allows for the services provided to be designed around community needs.

School-based health programs are costly; therefore, those who govern funding seldom mandate such programs until serious problems arise. Federal and state health care policy is then tapped to address the remedial/tertiary concerns. A critical issue is the degree to which programs should be targeted to specific at-risk groups using selective or indicated prevention approaches or spread across all groups with no differentiation using universal prevention approaches. Research has shown universal approaches to be far more likely to be successful, produce lasting change, and reduce long-term health costs. However, universal prevention programs may not be of sufficient dosage or targeted enough to have a discernible impact on higher risk children.

Most school-based centers must seek multiple financial supports. Obvious initial sources include the local school board and the local private sector. Several states have adopted laws or officially supported the development of school-based clinics. Federal or state block grants are often used to fund such programs. Medicaid may also be a source of funds if medically necessary services are provided to children and families who are Medicaid eligible. Managed care is another funding option. School-based clinics or school-linked health programs may become certified as providers for students whose parents are enrolled in a managed care plan. Finally, the CDC provides resources to ensure that the identified curriculum to reduce risk behaviors is available for state and local agencies interested in using it. The CDC directly funds coordinated school health programs in a number of states.

4. COORDINATED SCHOOL HEALTH PROGRAMS

A health-promoting school is a school in which all members of the school community work together to provide students with integrated and positive experiences and structures that promote and protect their health. This includes both formal and informal curricula in health, the creation of a safe and healthy school environment, the provision of appropriate health services, and the involvement of the family and the wider community in efforts to promote health.

Full-service schools (schools in which health, mental health, social, and/or family services may be located) or comprehensive health education are considered effective formats to promote healthy behaviors and prevent dysfunctional choices. However, without a radical change in school missions and society values (i.e., society choosing to directly address major issues such as poverty and unsafe environments), most of the 100,000 schools in the United States will not implement such a program in the near future. Selective or indicated prevention approaches are more consistent with current federal funding (need-based block grants) and perhaps offer the best option for the near future.

The School Health Programs has developed a comprehensive school health program model. According to the model, a comprehensive school plan provides

- A curriculum that addresses and integrates education about a range of categorical health issues at developmentally appropriate ages
- Programming for grades K–12
- Instruction for a prescribed amount of time at each grade level
- Integrative activities to help young people develop the skills they need to avoid the use of tobacco, dysfunctional behaviors, and engaging in a sedentary lifestyle
- Appropriate instruction from qualified teachers, health professionals, and concerned community members

Ideally, a school health-promotion model would be one that is holistic and that includes the interrelationships between the physical, social, emotional, and environmental (e.g., shaded play area) aspects of health. Health-promoting schools provide opportunities for families to take part in the development of health skills and knowledge of their children. Furthermore, an effective and comprehensive health-promotion model addresses the importance of the social ethos of the school in supporting a positive learning environment—an environment in which both healthy relationships and the emotional well-being of children are strengthened.

Schools, by themselves, should not be expected to address the nation's most serious health and social problems. However, school health programs can be one of the most efficient means of preventing major health and social problems. By reducing the incidence of mental and physical health problems that afflict young people and often continue into adulthood, school health programs may reduce the spiraling cost of health care, improve educational outcomes, and increase productivity and quality of life. Schools are widely recognized as the major setting in which activities should be undertaken to promote competence and prevent the

development of unhealthy behaviors. Furthermore, from a primary and secondary prevention perspective, an economic perspective, and a demographic standpoint, it appears essential that social initiatives such as health promotion be linked to schools.

4.1. Positive Youth Development

According to Larson, high rates of boredom, alienation, and disconnection from meaningful challenge are typically not signs of psychopathology but, rather, signs of deficiency of positive youth development. Furthermore, many cases of problem behavior, such as drug use, aggression, premature sexual activity, and minor delinquency, may also be related to an absence of a positive life trajectory. Although involvement in extracurricular activities, such as sports, art, and drama clubs, has been clearly demonstrated as beneficial in reducing high-risk behaviors and increasing academic success, these advantages are not being fully promoted by schools or engaged by a significant number of today's youth. Due to a variety of factors, many schools are not fully addressing the mental and other health needs of children and youth. Unfortunately, most efforts have been focused on when and how things go wrong for youth, while neglecting the positive aspects. Indeed, today's schools are so focused on the development of academic skills and raising test scores that health promotion and the benefits of extracurricular activities are largely ignored.

After-school programs offer an alternative that can aid in the promotion of positive youth development and provide a diversion between 3:00 and 8:00 PM, the peak hours for violent juvenile crimes. The development of "assets" has been key to a number of successful after-school programs. Survey data collected by the Search Institute identified 40 developmental assets that young people need to build in order to be healthy and successful. The survey data revealed the protective power of assets: The more assets a child has, the more likely it is that he or she will do well in school, engage in community service, and value diversity, and the less likely it is that the young person will engage in risk-taking behaviors such as drug or alcohol abuse, drop out of school, or engage in violent behavior. These assets are commonsense values and experiences that help youth to grow up caring, healthy, and responsible. Applying the assets as a framework for positive youth development provides schools and communities with a set of factors that are associated with fewer risk behaviors and increased health. Furthermore, the assets provide a common language for talking about what young people need and for pulling people together across different social groups.

Structured voluntary activities have promoted positive youth development and may be a key to successful in-school social and emotional learning. An adolescent's time is too often spent in two opposite experiential domains. In school, youth experience concentration and challenge without being intrinsically motivated. Conversely, in most leisure activities, they report experiencing intrinsic motivation but a lack of challenge and concentration. An exception to this is structured voluntary activities. Voluntary activities are those with rules, constraints, and goals as well as those in which the adolescent is highly motivated to be involved. They include sports, arts, hobbies, and participation in youth organizations. Based on a review of the literature, Larson concluded that involvement in structured voluntary activities is associated with multiple positive outcomes, including the development of initiative, greater achievement, and diminished delinquency. Thus, schools promote positive youth development when they provide opportunities for youth to engage in service learning and/or structured activities that contain the elements of intrinsic motivation, concentration, and challenge.

In recent years, a number of national reviews of school-based prevention programs have been performed to determine "best practices" in positive youth development. Using core competencies of social and emotional learning (SEL) as a basis for evaluation of programming and implementation, the Collaborative for Academic, Social, and Emotional Learning (CASEL) reviewed 81 programs that met its rigorous review standards. Following a 3-year review process, CASEL identified 22 programs that met its final criteria of outstanding SEL instruction, evidence of effectiveness, and outstanding professional development. Many of the programs selected were for all grade levels (K–12) because evidence-based/effective SEL programming typically begins at an early age and continues through high school.

5. SUMMARY

Today's youth face a plethora of risk factors and complex health challenges. In addition, the most serious and expensive health and social problems, as well as the leading causes of death for both youth and adults, result from negative behavioral patterns. Schools can be a dynamic setting for promoting health and an optimal entry point for enabling children to grow into healthy adults. Furthermore, health and education are

closely related: Good health provides children with better opportunities to pursue learning. Conversely, better education empowers children to engage in healthy practices and behaviors and act as change agents to promote the health of their families and communities. Thus, there are considerable gains to be made from partnerships between health and education.

A health-promoting school provides a setting in which education and health programs come together to foster health as an essential component of life. A health-promoting school responds to the health, educational, spiritual, and psychological needs of schoolchildren in an integrated and holistic way, in which interdisciplinary mechanisms are maximized. It is within this framework that schools can better assist students in becoming knowledgeable, responsible, and healthy citizens dedicated to the common good.

See Also the Following Articles

Alcohol Dependence ■ Bullying and Abuse on School Campuses ■ Emotional and Behavioral Problems, Students with ■ Forensic Mental Health Assessment ■ Health and Culture ■ Mental Measurement and Culture ■ Risk-Taking and Sexual Behavior ■ School Violence Prevention ■ Sexual Behavior ■ Sexual Behavior and Culture

Further Reading

Anderson, R. N. (2001). Deaths: Leading causes for 1999. *National Vital Statistics Report*, *49*(11), 1–88.

Annie E. Casey Foundation. (2003). *When teens have sex: Issues and trends*. Retrieved May 30, 2003, from http://www.aecf.org

Beginning too soon: Adolescent sexual behavior, pregnancy and parenthood. A review of research and interventions. (n.d.). Retrieved June 1, 2003, from http://aspe.os.dhhs.gov

Brenner, N. D., Martindale, J., & Weist, M. D. (2000). Mental health and social services: Results from the school health policies and programs study 2000. *Journal of School Health*, *7*(7), 305–312.

Center for Health and Healthcare in School. (n.d.). *Children's mental health needs and school-based services: A fact sheet*. Retrieved December 5, 2002, from http://www.healthinschools.org

Centers for Disease Control and Prevention. (2001). *Youth risk behavior surveillance: United States 2001*. Retrieved December 7, 2002, from hhwww.cdc.gov

Collaborative for Academic, Social, and Emotional Learning. (2003). *Safe and sound: An educational leader's guide to evidence-based social and emotional learning programs*. Chicago: Collaborative for Academic, Social, and Emotional Learning.

Focus Adolescent Services. (2001). *Ecstasy (MDMA) and club drugs*. Retrieved May 30, 2003, from http://www.focusas.com

Larson, R. W. (2000). Toward a psychology of positive youth development. *American Psychologist, 52,* 256–265.

McQueen, A. (1998). *Study links teen promiscuity with drugs, alcohol*. Retrieved June 1, 2003, from http://cyi-stars.org

National Youth Violence Prevention Resource Center. (2003). *Teens and alcohol*. Retrieved May 30, 2003, from http://www.safeyouth.org

Partnership for a Drug-Free America. (2003). *Partnership attitude tracking study*. Retrieved June 1, 2003, from http://www.drugfreeamerica.org

Payton, J. W., Wardlaw, D. M., Graczyk, P. A., Bloodworth, M. R., Tompsett, C. J., & Weissberg, R. P. (2000). Social and emotional learning: A framework for promoting mental health and reducing risk behaviors in children and youth. *Journal of School Health*, *70*(5), 179–185.

RAND. (2001). *Health research highlights: Mental health care for youth*. Retrieved December 4, 2002, from http://www.rand.org

Rones, M., & Hoagwood, K. (2000). School-based mental health services: A research review. *Clinical Child & Family Psychology Review*, *3*(4), 223–241.

Satcher, D. (1999). Getting physical: Exercising our demons: Sedentary lifestyles and fatty foods leave Americans overweight and unhealthy. *Atlanta Journal Constitution,* p. A17.

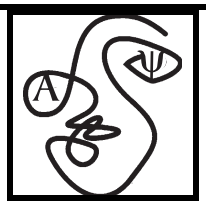
Saunders, R., & Pate, R. (2000). Promoting physical health. In K. M. Minke, & G. G. Bear (Eds.), *Preventing school problems—Promoting school success*. Bethesda, MD: National Association of School Psychologists.

School Health Programs. (2000). *Center for Disease Control and Prevention at-a-glance 2000*. Retrieved December 7, 2002, from http://www.cdc.gov

Search Institute. (2000). *Take it to heart: An asset-based guide to nurturing children, youth, and families in faith community* [Brochure]. Minneapolis, MN: Author.

U.S. Department of Health and Human Services. (2001). *America's children: Key national indicators of well-being, 2001*. Retrieved December 12, 2002, from http://www.childstats.gov

Zins, J. E., & Wagner, D. L. (1997). Health promotion. In G. G. Bear, K. M. Minke, & A. Thomas (Eds.), *Children's needs II: Development, problems, and alternatives* (pp. 945–954). Bethesda, MD: National Association of School Psychologists.



Health Psychology, Cross-Cultural

Malcolm MacLachlan
Trinity College Dublin, Dublin, Ireland

Mesfin Samuel Mulatu
National Institute of Mental Health, Bethesda, Maryland, USA

1. Concepts and Definitions
2. Theoretical Foundations
3. Cross-Cultural Health Psychology in Practice
4. Current Issues and Future Directions
Further Reading

GLOSSARY

acculturation A process that occurs when people from one culture encounter and react to people from another culture; it may occur at an individual (psychological) level or at a group (social) level.

cross-cultural psychology A specialized field in psychology with an approach that compares and contrasts individual psychological functioning in various cultural contexts.

cultural psychology A specialized field in psychology with an approach that roots behavior in its cultural context; it views culture and psychology as "making each other up."

ecocultural framework An approach that seeks to understanding cultural similarities and differences as adaptations to various human and environmental contexts.

global health An attempt to address health problems that transcend national boundaries, may be influenced by circumstances and experiences in other countries, and are best addressed by cooperative actions and solutions; the underlying assumption is that the world's health problems are shared problems and, therefore, are best tackled by shared solutions.

health inequality A term that refers to disparities or variations in health status achieved by individuals, groups, or communities; in its frequent form of use, it refers to differences in the levels of health between men and women, between majority and minority communities, and between the rich and poor within societies.

health psychology A specialized field in psychology that studies the psychosocial aspects of mental and physical health and illness as well as the contexts in which these occur.

HIV/AIDS Abbreviation for human immunodeficiency virus/acquired immune deficiency syndrome; HIV is a retrovirus that infects cells of the immune system (primarily the CD4 cells and macrophages) and then destroys or impairs their functions, leading to immune deficiency and resulting in a disease complex known as AIDS.

primary prevention Health activities that are designed to reduce or eliminate the risk of acquiring a health problem; for example, interventions that encourage sexual abstinence or avoidance of intravenous drug use represent primary prevention efforts against HIV infection.

secondary prevention Health activities that focus on minimizing or eliminating the impact of an already acquired health problem on the affected individuals, families, and communities or on slowing or stopping the progress of a disease during its early stages; an example among AIDS patients may include efforts aimed at preventing opportunistic infections and at managing pain and psychosocial distress.

Cross-cultural health psychology is defined by examining what each of the components of this term means in its own right and by acknowledging the challenges that arise by combining them. Cross-cultural health psychology incorporates aspects of health psychology, anthropology, behavioral medicine, and public health. It is primarily concerned with

understanding health and illness from a psychological perspective, but within the context of its cultural constructions. This article reviews some relevant models that combine these individual and community perspectives. For illustrative purposes, from the vast array of health problems that are addressed by cross-cultural health psychology, the article focuses its review on HIV/AIDS due to its pandemic nature. The article considers risk factors, determinants, and prevention; disease management and impact alleviation; and social inequalities and, more broadly, impacts on population health. It concludes by arguing for a broadening of the conceptualization of what "cross-cultural" means and notes seven themes that relate culture to health. The effects of globalization and the increasing complexities of what constitutes cross-cultural health psychology present the latter with an opportunity to make a valuable contribution to the emerging field of global health.

1. CONCEPTS AND DEFINITIONS

In defining cross-cultural health psychology, one needs to do so reflexively by unpacking each of the words in the term. If one takes "psychology" as the study of human behavior, one must qualify this further by acknowledging that "psychology" has mostly been understood to be synonymous with a particular perspective on human behavior. This perspective is characteristic of a broader rationalist, reductionist, and individualist approach to advancing knowledge that predominates in Northern and Western conceptions of what are valid means of truth seeking. This is problematic for psychology because such an approach is, in itself, a cultural construction. However, other psychologies, curiously referred to as "indigenous," relate different conceptions of how one ought to account for human behavior. Thus, one first needs to acknowledge that there are different—culturally constructed—conceptions of just what psychology is and how it should study human behavior.

In some ways, "health" is easier to define, if only because the World Health Organization's (WHO) definition is rather broad and inclusive in emphasizing biological, psychological, and social aspects of health and in acknowledging that health and illness are not "all or none" categories. All people are healthy or ill to some extent, and even those who suffer greatly through physically disabling diseases may be quite healthy in other respects, whereas some icons of physical excellence and beauty may be quite ill in some ways. Of course, the way in which distress is expressed may differ among cultural contexts (e.g., by psychologizing or somatizing), as may the understanding of what it means to be healthy or ill and the ways in which to respond to illness.

Therefore, "health psychology" may be taken to be the study of how human behavior influences health, broadly defined. "Health" here refers to well-being in general and is not confined to either physical complaints or mental complaints. Health psychology seeks to bridge the Cartesian divide by acknowledging that human experience is mediated through the body and, thus, is embodied. The biopsychosocial model, so often identified with health psychology, should be seen not as multiple but rather as integrative, that is, as being of one—body and mind together. Furthermore, the contextual aspects of health, in terms of its social settings and conditions, should also be to the fore in health psychology.

"Cross-cultural" psychology must always first be "cultural" psychology for it to be meaningful. Cultural psychology seeks to understand how behavior is influenced by the social context in which it occurs. It further acknowledges that this context is woven through particular customs, rituals, beliefs, ways of understanding and communicating, and so on, so that distinctive patterns of behavior are cultivated. Only by understanding how a culture patterns meaning can we be sure to know that the sort of things we might want to compare among different cultures actually have some similarity. Having established a meaningful similarity in the structure or function of aspects of human behavior, it may then be enlightening to compare such behavior in different cultural contexts, that is, across cultures. Cross-cultural psychology may incorporate both the localized cultural perspective and the comparative cultural perspective.

In sum, cross-cultural health psychology seeks to describe the diversity of health (and illness) behavior and to link this to the cultural context in which it occurs. It recognizes that culture and psyche "make each other up" and that the ways in which they do this have implications for health. The value of cross-cultural health psychology is becoming increasingly recognized not only in the treatment of health problems that present in specific cultural contexts but also in how social and cultural factors must be understood for the prevention of illness and the promotion of health.

2. THEORETICAL FOUNDATIONS

Cross-cultural health psychology borrows theoretical and practical approaches from a wide variety of related disciplines, including the broader fields of psychology, anthropology, behavioral medicine, and public health. Health psychology and behavioral medicine contribute the biopsychosocial model of health and disease. A recently revised version of this model argues for the reorientation of the fields of health and medical sciences such that research and clinical practice encompass both health and illness. The model aims to decipher the interactive and reciprocal relationships among biophysical, psychological, and social processes in health and disease. It takes into account the social networks of the individual to achieve, maintain, and maximize health and well-being for individuals, their significant others, and society. The major challenge for this model is for its biological, psychological, and social aspects to truly act in synthesis rather than as simply parallel and complementary aspects of health and illness.

Cross-cultural psychologists, anthropologists, and others argue that current theoretical models and practice in medicine and health psychology are insufficient to capture the centrality of culture to the human experiences of health and illnesses. They argue for theoretical frameworks that recognize the ubiquitous effect of culture on health and illness; that treat culture as a system of meaning that links body, mind, and world; and that promotes the significance of cultural context as a resource. Frameworks that give greater weight to cultural and ecological contexts and process as they relate to health and health promotion have been suggested. For instance, Berry's ecocultural framework provides a scheme with which factors at both the individual (psychological) and community (cultural) levels are considered in cross-cultural application of research in cognitive, affective, behavioral, and social aspects of health phenomena. The community level of work typically involves ethnographic methods and yields a general characterization of shared health conceptions, values, practices, and institutions in a society. The individual level of work involves the psychological study of a sample of individuals from the society and yields information about individual differences (and similarities) that can lead to inferences about the psychological underpinnings of individual health beliefs, attitudes, behaviors, and relationships. There are important reasons why cultural- and individual-level factors need to be taken into account when dealing with health phenomena. Cultural-level health phenomena are taken into account to establish the broad context for the development and display of individual health phenomena. Without an understanding of this background context, attempts to deal with individuals and their health problems may well be useless or even harmful; individuals are not "culture free." The reason for considering individual-level health phenomena is that not all persons in a cultural group hold the same beliefs or attitudes, nor do they engage in the same behaviors and relationships. Without an understanding of people's individual variations from the general community situation, harm may well be inflicted again; individuals are not "cultural dopes."

Health models that take into account broader cultural- and ecological-level factors have also been proposed in the field of public health. For instance, Airhihenbuwa proposed a theoretical model that places cultural and environmental factors at the core of the development, implementation, and evaluation of health programs. His PEN-3 model consists of three dimensions of health beliefs and behaviors: (a) health education, (b) educational diagnosis of behavior, and (c) cultural appropriateness of health behavior. Each of these three interdependent dimensions is influenced by three categories of factors (representing the acronym "PEN"). Accordingly, health education needs to be designed to address the health needs of all, involving the person, the extended family, and the neighborhood or community in which the person lives. Airhihenbuwa noted that educational diagnosis of health behavior helps researchers and program developers to determine the factors that influence individual, family, and/or neighborhood and community health actions. Critical elements that need to be assessed as part of such a diagnosis include perceptions, enablers, and nurturers. Health-related perceptions include attitudes, values, and beliefs within a cultural context that facilitate or hinder health-promoting or disease-preventing actions. Enablers are distal cultural, societal, systemic, or structural factors that may enhance or be barriers to health actions, whereas nurturers are relatively proximal influences on health behaviors originating from extended family, kin, friends, and the neighborhood. Finally, the most critical dimension of developing effective health programs is the determination of the cultural appropriateness of the health behavior. Within a given cultural context, some health beliefs and behaviors can be positive, endowing the person, the family, and/or the community with known health benefits. Other health beliefs and behaviors may be existential, having no harmful health consequences

but bearing important indigenous meaning to the community. Still others may be negative health beliefs and behaviors, contributing to the origin or exacerbation of health problems. Such an assessment of the cultural appropriateness of health beliefs and behaviors empowers health program developers to consider the cultural perspectives, knowledge base, and resources of the people for whom the programs are designed.

A relatively more extensive framework for planning health promotion programs is the PRECEDE/PROCEED model developed by Green and Kreuter. This model takes into account the myriad of factors that determine the health and quality of life of individuals and communities in planning, implementing, and evaluating health programs. The PRECEDE part of the model (the acronym stands for predisposing, reinforcing, and enabling constructs in educational/ecological diagnosis and evaluation) involves various phases during which the determinants of health are diagnosed. It begins with a social assessment to identify quality of life and concerns of the target population, followed by assessment of health problems and behavioral and environmental factors linked to these health problems. Then, the predisposing, reinforcing, and enabling factors that influence health behaviors are assessed during the educational and ecological assessment phase, followed by an assessment of the organizational and administrative capabilities and resources for the development and implementation of health programs to address the problems identified. The PROCEED part of the model (the acronym stands for policy, regulatory, and organizational constructs in educational and environmental development) offers additional steps for developing policy and initiating the implementation and evaluation process appropriate to the health outcomes, intervention types, and target populations. The PRECEDE/PROCEED model, like PEN-3, recognizes that health is determined by multiple factors and that interventions need to take multilevel and multidimensional actions to be effective.

More specific individual-level health behavior theoretical models (e.g., Health Belief Model, theory of planned behavior, social cognitive theory) are often used in health psychology and public health to explain and predict risk and protective behaviors. Embedded in these models is the belief that risk behaviors, cognitions, and their determinants are identifiable, modifiable, and causally linked to health outcomes. In general, these health behavior models propose that one or a combination of the following factors predict the adoption and maintenance of changes in health behaviors: (a) knowledge and information about the health threat; (b) perceived severity of and personal vulnerability to the threat; (c) perceived control or self-efficacy to adopt the behavior change; (d) perceived efficacy of the behavior change to prevent the threat; (e) perceived benefits compared with costs associated with behavioral change; (f) intentions, attitudes, and social norms supporting the adoption of the behavioral change; (g) actual skills to bring about the behavioral change; and (h) availability of cues for behavioral change. Recognizing the limitation of individual-focused factors to behavior change, other theoretical models in public health and related disciplines have expanded this list by adding variables at the interpersonal (e.g., couple communication, power in relationships), community (e.g., supportive norms, social networks), and environmental and structural (e.g., accessibility services, policies) levels that have strong bearings on individual-level risk behaviors and cognitions.

In sum, combining applied individual- and community-level methodologies from clinical psychology and public health and the conceptual underpinnings from interactive biopsychosocial and ecocultural frameworks, cross-cultural health psychology focuses on factors that determine and maintain health and health outcomes in cross-cultural and multinational contexts. This approach presupposes consideration of multiple causal pathways and levels of analysis in research and practice of health psychology. Psychosocial, cultural, and environmental factors are believed to be causally linked to health directly through exposure to noxious and disease-causing conditions such as carcinogens and pathogens. Indirectly, psychosocial stress associated with adverse psychosocial, cultural, and environmental conditions influences health by affecting the performance of health-related behaviors as well as the functioning of the central nervous system and the endocrine system, primarily immune and cardiovascular functioning. Overall, it is believed that biological, psychological, sociocultural, and environmental processes interactively affect, and are affected by, health and illness.

3. CROSS-CULTURAL HEALTH PSYCHOLOGY IN PRACTICE

Applied health psychology consists of several research and practice areas that overlap with the broader fields of clinical psychology, behavioral medicine, and public health. In cross-cultural contexts, three areas of research and practice have been particularly predominant: (a) identifying health risk factors and their determinants

and taking preventive actions, (b) alleviating and managing the impacts of diseases, and (c) addressing social inequalities and population health. Although a vast range of health problems could be considered here, this section discusses each of these three areas by giving pertinent examples from the literature on human immunodeficiency virus/acquired immune deficiency syndrome (HIV/AIDS) due to its pandemic nature. Extensive coverage of cross-cultural applications of health psychology in understanding and tackling other major global health problems, including malaria, tuberculosis, nutritional deficiencies, and noncommunicable diseases, can be found elsewhere.

3.1. Risk Factors, Determinants, and Prevention

Before focusing exclusively on HIV/AIDS-related issues, the first two paragraphs of this section briefly consider the spectrum of disease and mortality globally. According to *The World Health Report 2002* published by the WHO, the leading causes of death in high-mortality "developing" societies (e.g., Sub-Saharan Africa, Latin America, Asia) are communicable diseases and maternal and child conditions, including HIV/AIDS, tuberculosis, malaria, respiratory infections, and childhood and maternal undernutrition. In highly "developed" countries (e.g., North America, Western Europe, Pacific Rim), the leading causes of death are noncommunicable diseases, including various forms of cardiovascular diseases, cancers, and psychiatric disorders. In low-mortality developing countries (e.g., Middle East, Eastern Europe, Southeast Asia), a mix of communicable and noncommunicable diseases and maternal and childhood conditions characterize the leading causes of death.

These differences in leading causes of death among countries and regions have been attributed to differences in risk factors to which individuals in these societies are exposed. This is a subject of interest in cross-cultural health psychology research and practice. According to the WHO report, in the high-mortality developing societies, childhood and maternal undernutrition, unsafe sexual activities, indoor smoke from solid fuels, and unsafe water, sanitation, and hygiene contribute the most to the total burden of ill health. In contrast, in highly developed countries, tobacco, high blood pressure, alcohol, high cholesterol, and high body mass index play major roles in the loss of health and functioning. In low-mortality developing countries, a mixed pattern of risk factors found in both high-mortality developing and low-mortality developed countries seem to contribute significantly to disability and death. Two important themes appear to characterize this report. First, cross-cultural and cross-national differences in health and disease are associated with socioeconomic development and geopolitical situations of the countries and regions. Second, the single most important causes of burden of disease and disability in both developing and developed countries are modifiable health behaviors or conditions that are responsive to appropriate health measures.

Focusing on HIV/AIDS, the most extensively studied risk factor is unprotected sexual intercourse. With the rapid spread of HIV/AIDS, other sexually transmitted infections, and other reproductive health problems, increased attention has been given to understanding the patterns and determinants of sexual risk behaviors and to designing and implementing effective interventions. Although HIV as the cause of AIDS and risk factors for its transmission (e.g., unprotected sex, contaminated blood transfusion, mother-to-child transmission, unsterilized injections) have been known since the early 1980s, controlling its spread has been a formidable challenge to public health across the world, particularly in Sub-Sahara African countries where the epidemic has hit the hardest.

Health psychologists working in a cross-cultural context have applied health behavior and social psychological theories in both research and applied settings. In research, health psychologists are involved in the design and implementation of rapid assessment techniques to measure the level of knowledge, attitudes, and practices of individuals in general or of those who are considered to be particularly at risk for infection. These efforts have become expanded and standardized and are being used in multiple national surveys on a regular basis with the intent to collect comparable risk and protective data. An important example in these efforts is the Demographic and Health Surveys, a worldwide initiative supported by the U.S. Agency for International Development (USAID) and governments of several developing countries with the aim of providing comparable data and analysis on population health, including HIV/AIDS. These studies, and similar other studies, help to determine the levels of awareness and attitudes about HIV/AIDS and the extent of risk behaviors for HIV infection. Based on these data, it is known that awareness about HIV/AIDS and its primary routes of infection is high in nearly every country but that high-risk behaviors continue to be practiced to varying degrees in both developed and developing countries. For example, unprotected heterosexual intercourse is a significant means of HIV spread in

Sub-Saharan Africa, whereas unprotected homosexual intercourse and injecting drug use are major sources of infection in Western countries.

Research has also focused on the determinants of sexual risk (e.g., sex with multiple partners) or protective behaviors (e.g., condom use) using general health behavior theories (e.g., theory of planned behavior) or those that are developed specifically to address HIV/AIDS-related risk (e.g., AIDS Risk Reduction Model). According to a recent report on the epidemic by the UN Program on HIV/AIDS (UNAIDS), sexual risk taking or inconsistent condom use is closely intertwined with a variety of personal and interpersonal psychosocial factors, including poor knowledge, a low sense of threat to infection, high sensation seeking, and a strong need to comply with social norms, as well as with group- or community-level risk factors, including poverty, migration, and pervasive norms that perpetuate women's subordination within sexual relationships. In the context of Sub-Saharan Africa, negative attitudes toward sex with condoms, female sexual compliance to ensure economic security, a lack of negotiation skills to implement condom use, a desire for children, and an external locus of control have been identified as potential obstacles to condom acceptance and use. The relative strength of individual-, interpersonal-, community-, and structural-level factors varies from society to society, depending on the level of socioeconomic development, the spread of the disease, and the sociopolitical climate of the country. However, it is now recognized that sexual risk and protective behaviors, like other health behaviors, are determined by multiple factors and imply that interventions need to address many factors to be successful.

Interventions against risky sexual behaviors often teach recipients about facts related to HIV and reproductive biology, encourage positive attitudes toward those affected and infected by HIV/AIDS, instill skills in sexual negotiation and in the acquisition and use of condoms, and make condoms and voluntary testing and counseling more accessible. Given that sexual beliefs, attitudes, and activities are strongly influenced by cultural and broader structural factors, some interventions have addressed these factors appropriately. Most interventions include a combination of behavioral outcomes, theoretical approaches, and formats and methods of delivery. In some countries and regions, some formats are more acceptable and amenable than others. For instance, in Sub-Saharan Africa, radio and stage dramas are often used for communicating preventive messages. In more literate societies, print and other mass communication approaches are used.

Evidence from both more and less industrialized countries indicates that behavioral and psychosocial interventions are effective in enabling people to make and maintain necessary preventive sexual behavior changes. Reviews of these interventions generally conclude that psychosocial and behavioral interventions are effective in increasing awareness, skills in sexual negotiation, frequency of communications with sexual partners about safer sex, and skills in the acquisition and use of condoms as well as in reducing frequency of sexual intercourse, early sexual initiation among youth, and the incidence of sexually transmitted infections. Although intervention effectiveness varies by the type, length, and other characteristics of the intervention, the most efficacious HIV prevention programs have been directed specifically at groups at risk for infection, focused on relationship and negotiation skills, involved multiple sustained contacts, and used a combination of culturally appropriate media of delivery. Although most intervention studies have not explicitly studied the role that cultural variables play within cross-cultural contexts, those that have integrated cultural factors into their HIV prevention themes and contents have confirmed the pervasive influence of culture on risk perception, risk-taking behaviors, and adoption of protective behaviors.

3.2. Disease Management and Impact Alleviation

Advances in health sciences, care, and technology have created important new challenges and roles for cross-cultural health psychologists in the area of secondary prevention. Many of the acute and chronic medical problems facing the world, and the curative or palliative treatments for these problems, have profound impacts on numerous areas of patient functioning and quality of life. These impacts may involve physical and psychosocial disabilities and may originate directly from the health problems themselves, from treatment side effects, or indirectly as a consequence of the disabilities or limitations they impose on patients and families. In general, medical conditions present psychosocial challenges for patients and their families in the areas of (a) coping with and managing pain, (b) coping with psychosocial impacts of disease, (c) adhering to demanding medical regimens, and (d) coping with social stigma and discrimination.

Chronic health problems, such as HIV/AIDS, are accompanied by persistent pain and other psychosocial difficulties in the form of emotional distress, including

anxiety, depression, low self-esteem, and irritability, as well as disruptions in interpersonal relationships, including marital difficulties and family stress. Pain and other psychosocial consequences of diseases must be addressed not only because they contribute to diminishing quality of life of the patients and their families but also because studies have documented that psychosocial distress is related to length of survival, immunological functioning, and/or some measures of disease progression. Sources of persistent pain and psychosocial distress among people with chronic diseases are multiple. Among people living with HIV/AIDS, psychosocial problems occur as a consequence of the physical illness, disability, and wastage and also as side effects of the multitudes of medication taken. Neurological and cognitive impacts of HIV infection may also affect the day-to-day functioning of patients. Negative attitudes and reactions of others toward patients contribute to patients' sense of isolation and psychosocial distress. Furthermore, socioeconomic strains caused by the loss of jobs and incomes add to the psychosocial distress that the patients and their families experience. The actual level of psychosocial distress experienced by people living with HIV/AIDS depends on several factors, including the resources available to patients and their families, socioeconomic development of their countries, and the strength of national policies protecting the rights of HIV-infected people. For instance, the potential loss of jobs, housing, family support, and even life due to admitting to HIV infection is a major concern and source of distress in many developing societies, although such concerns are significantly reduced in Western societies.

Interventions against persistent pain and psychosocial distress due to HIV/AIDS involve multidisciplinary teams and a variety of components and strategies, including individual psychotherapy, social support in a group setting, cognitive–behavioral skills training, and pharmacological treatment to reduce and manage pain, anxiety, and depressive symptoms. Self-management interventions are also used when the aim is to help patients monitor and cope with their pain and illness and to maintain satisfactory quality of life. These interventions include providing information about treatment conditions, management of symptoms and psychological consequences, and communication and coping strategies such as relaxation and meditation.

Reviews of studies from various countries and using various outcome measures have found strong support for interventions against persistent pain and psychosocial consequences of HIV/AIDS. Although the positive effects of these interventions are not always consistent, and it is difficult to discern the relative efficacy of their components, there is sufficient evidence that interventions against psychosocial consequences of diseases can affect quality of life as well as disease outcomes. Consistent with the biopsychosocial and ecocultural perspectives of both pain and psychosocial distress, there is strong evidence that culturally appropriate interventions that take into account sociocultural resources, explanatory models, and contextual factors are more likely to effectively reduce persistent pain, psychosocial distress, and negative disease outcomes. For instance, the use of traditional healers, midwives, and indigenous social support systems in developing countries and the use of "alternative and complementary therapies" in the West have been found to be potentially useful interventions for managing pain and for mitigating the psychosocial impacts of HIV/AIDS.

Low adherence to prescribed treatments is a ubiquitous problem among patients with chronic conditions, thereby contributing to undermining the potential benefits that may accrue from treatment. Adherence to highly active antiretroviral therapy (HAART) and other anti-HIV medication has been studied in various cultural settings. In these studies, health psychologists are involved in examining the reasons for nonadherence and in designing interventions to decrease its likelihood. Studies from several countries and cultures have identified several factors that are correlated with nonadherence. Poor adherence tends to occur in treatment regimens that are complex, have strong side effects, and place too many cognitive and scheduling demands. Psychosocial factors, including emotional distress, low self-efficacy beliefs, substance abuse, pessimism about prognosis, lack of social or provider support, and limited access to health service resources, have also predicted lower levels of adherence. Thus, both individual-level variables and cultural constructions regarding the meaning of AIDS are important regarding adherence to treatment regimens.

Interventions against nonadherence with anti-HIV medication involve a combination of cognitive, behavioral, and psychosocial components. Health psychologists, along with other health care professionals, provide structured dosing instructions, self-monitoring and problem-solving skills, behavioral reminders and strategies to increase compliance with medical advice, counseling and social support in either individual or group settings, and information about drug options, interactions, and side effects. Studies that evaluated these interventions showed that patients who participated in interventions had higher levels of adherence and lower levels of viral

load than did patients who took standard care, although these findings are not always statistically significant and long-lasting. In addition, it has been noted that comprehensive interventions that combine educational, behavioral, and affective components have led to better intervention outcomes than have single-focused interventions. Effective strategies are also characterized by effective communications with patients that are tailored to individual patients' needs, capacities, and resources.

Stigma refers to undesirable or discrediting socially constructed attributes that individuals possess and that affect their quality of life as well as their social standing. Perceived or felt stigma, as opposed to enacted stigma or discrimination, involves real or imagined fear of negative social attitudes and discrimination due to one's undesirable personal attributes or association with others who have such attributes. Stigma is a powerful force that often causes prejudicial thoughts, behaviors, and actions by individuals, groups, communities, and institutions. Although nearly all health conditions can lead to stigmatizing experiences, diseases that elicit fear of illness, contagion, and death, as well as those whose origins are uncertain and whose treatments are limited, tend to cause a relatively greater amount of stigmatization, Patients with diseases such as leprosy and tuberculosis have historically evoked fear and revulsion in others. Although advances in knowledge and treatment have increased acceptance of such patients, other more recent health conditions, such as HIV/AIDS, are currently among the most stigmatizing health conditions around the world.

Among HIV-infected individuals, stigma manifests itself in many ways. HIV-infected individuals may refuse to disclose their status to others due to fear of negative reactions from their partners, family members, employers, and community members. This has been found to be a major obstacle to preventive efforts in many Sub-Sahara African countries. These individuals may also refuse to receive the necessary health services due to potential negative reactions of health care providers or to fear of being discriminated against by others. HIV-infected people who belong to certain sectors of society, such as commercial sex workers in Sub-Saharan Africa and injecting drug users in Western countries, may face an extra layer of stigmatization due to their association with "deviant" or "immoral" behaviors. These and other stigmatizing experiences lead to negative effects on test-seeking behavior, willingness to disclose HIV status, health service-seeking behavior, quality of health care received, and social support solicited and received.

Reviews of intervention studies from the United States, Europe, and Sub-Saharan Africa indicate that antistigma efforts aim at increasing the tolerance of individuals and communities for people living with HIV/AIDS, the willingness of health care providers to treat people with HIV/AIDS, and the abilities of people living with HIV/AIDS to cope with stigma and discrimination. Interventions have used a combination of strategies, including providing information about HIV/AIDS and impacts of stigma, building skills to diffuse and cope effectively with conflict situations, counseling to address individual concerns and provide social support, and providing opportunities for public contact with people infected with, and affected by, HIV/AIDS. The review indicates that interventions that combine information, skill building, and contact with people living with HIV/AIDS reduce stigma significantly more than do interventions that include any one of these separately. Not all studies have brought appreciable reductions in stigma, due partly to limitations in the conceptualization, measurement, and methodological problems associated with the interventions.

3.3. Addressing Social Inequalities and Population Health

Recently, research and practice in social inequality in health has gained ground in health psychology and public health across nations. There is a growing awareness that socioeconomic hierarchies, defined by a wide variety of indexes such as social class, accumulated wealth, occupational prestige, income, and education, and negative social attitudes and actions associated with those hierarchies, including racism, discrimination, and stigmatization, have been linked to differences in levels of health outcomes. Major national and multinational studies indicate that unfavorable outcomes in many areas of health are relatively more prevalent among people occupying the lowest socioeconomic status and/or among historically marginalized ethnic groups and immigrants. For instance, African and Hispanic Americans in the United States, Bangladeshis and Pakistanis in the United Kingdom, and Africans and Asians in South Africa tend to have unfavorable socioeconomic standing as well as poorer overall health than do White people in their respective countries. Health psychologists, as well as other professionals, are involved in efforts directed at understanding the mechanisms through which these effects occur and the manner in which these social inequalities in health can be reversed.

Research on ethnic densities has found that the larger the number of people in a minority cultural group living within a given area, the lower the psychiatric admission rate for members of that group, suggesting that living in a neighborhood that encourages an unfamiliar lifestyle, culture, and language is a risk factor for the development of mental disorder. Such segregation may help to protect group members from direct prejudice and provide them with social support that operates through culturally familiar customs. Research into the broader mechanisms of socioeconomic differences in health attributes their origins to two broad factors: material deprivation and psychosocial disadvantage. Included in the material explanation of social inequalities in health are access to basic necessities of life (e.g., food, shelter), access to amenities that contribute to a comfortable and productive life (e.g., home ownership, means of transportation), and access to curative, preventive, and/or health-promoting services (e.g., health care, exercise facilities). The psychosocial interpretation of health inequalities proposes direct and indirect effects of psychosocial stress associated with occupying lower socioeconomic status. Stress and emotional distress associated with racism, discrimination, and living in an unfavorable social position may relate to health directly by affecting physiological systems or indirectly by affecting health behaviors.

Efforts to address social inequalities in health and to promote population health involve a wide array of interventions to increase disease-preventing/health-promoting resources of individuals, strengthen capacities of communities, improve access to essential facilities and services, and encourage macroeconomic and sociocultural changes by way of conducive policies and government actions. A strong argument in favor of bringing changes in social policies and government programs that directly address the fundamental social causes of social inequalities in health is critical. For instance, structural-level interventions focusing on increasing accessibility of HIV/AIDS preventive and treatment services and on enacting favorable national policies have shown considerable potential in addressing the disproportionate burden of HIV/AIDS among ethnic and racial minorities and vulnerable groups in both resource-poor and resource-rich countries. Cross-cultural health psychology contributes to knowledge about behavior change approaches that incorporate culturally appropriate strategies. It can also bring forth knowledge about the social world and community-level variables (e.g., supportive social networks, social cohesion) that contribute to resilience, wellness, and health. Social and health policies that are evidence based and socioecologically appropriate are more likely to be effective in reducing social inequalities and improving individual and population health.

4. CURRENT ISSUES AND FUTURE DIRECTIONS

Traditionally, cross-cultural health psychology has sought to compare health in various cultural contexts, where those contexts either were geographically distant or involved migration from one geographical area into the culture of another geographical area. Although these fields of study are still of interest and relevance, perhaps a broader conceptualization of what "cross-cultural" means is required. Furthermore, it is important to recognize that cultural comparisons are naive if they fail to take account of the broader sociopolitical and economic context. This is particularly relevant to health when cultural differences are compounded by the disempowerment of one group relative to another. For instance, in thinking of the practical implementation of health promotion initiatives in Sub-Saharan Africa, it is clear that traditional indigenous approaches to health are not operating on an equal power base with more Western-oriented biomedical interventions. Aid agencies that provide resources for health care tend to support those perspectives that reflect their own cultural assumptions. Although this is not necessarily wrong, cross-cultural health psychology must be aware of this bias and power imbalance to be truly informative for effective action.

MacLachlan identified seven realms where culture and health are related to each other and where these relationships have implications for people's sense of identity, empowerment, and well-being (Table I). These different realms reflect a range of new challenges for cross-cultural health psychology, where "cross-cultural" may be taken to mean between different realms of culture. Hermans and Kempen argued that cross-cultural psychology's traditional emphasis on geographical location is becoming increasingly irrelevant. Instead, they emphasized the influence of interconnecting cultural and economic systems, of cultural hybridization creating multiple identities, and of increasingly permeable cultural boundaries. Most contemporary societies are patterned by a multiplicity of meanings and practices that likely contribute to greater uncertainty. As Hermans and Kempen stated, "Uncertainty is not primarily in a culture's core but in its contact zones."

Table I
Realms within Which Culture and Health Interact

Cultural colonialism	Rooted in the 19th century, when Europeans sought to compare a God-given superior "us" with an inferior "them" and to determine the most advantageous way of managing "them" in order to further European elites.
Cultural sensitivity	Being aware of the minorities among "us" and seeking to make the benefits enjoyed by mainstream society more accessible and modifiable for "them."
Cultural migration	Taking account of how the difficulties of adapting to a new culture influence the opportunities and well-being of geographical migrants.
Cultural alternativism	Different approaches to healthcare offer people alternative ways of being understood and of understanding their own experiences.
Cultural empowerment	As many problems are associated with the marginalization and oppression of minority groups, a process of cultural reawakening offers a form of increasing self and community respect.
Cultural globalization	Increasing (primarily) North American political, economic, and corporate power reduces local uniqueness and reinforces and creates systems of exploitation and dependency among the poor, throughout the world.
Cultural evolution	As social values change within cultures, adaptation and identity can become problematic, with familiar support systems diminishing and cherished goals being replaced by alternatives.

After MacLachlan (2004b).

Conventionally, cross-cultural health psychology has recognized the interplay between culture and health in colonialism, through sensitivity to minority groups, and in the acculturation issues of migrants. The recent enthusiasm for alternative medicine has also attracted the interest of some cross-cultural health psychologists because it can be seen as introducing greater pluralism into health care. However, less noticed by cross-cultural health psychologists are the distinctive (and often more complex) relationships between culture and health that are evident in the renewed enthusiasm for cultural empowerment, especially among marginalized indigenous groups. Another complex issue is the health effects of globalization, where a singular politicoeconomic system based on North American values, with accompanying ontological assumptions about the causes of and cures for health problems, is in conflict with local practices and beliefs. Yet another challenge for cross-cultural health psychology is how rapid social and cultural changes within societies may bring in their wake what might be termed "disorders of acculturation." For instance, cultural evolution and social change may require a process of temporal acculturation that has significant implication for health.

Although the realms described here may require cross-cultural health psychologists to rethink some of their discipline's traditional assumptions, some aspects of traditional health psychology will be applicable in the "newer" realms where culture and health interplay. For instance, Berry's framework for conceptualizing the acculturation strategies of (geographical) migrants has also been found to have predictive value for the well-being of people experiencing temporal acculturation through a period of rapid social change. There is no doubt that cross-cultural health psychology offers a valuable comparative perspective on the human experience of health across cultures. As an orientation, it has much to offer. As a methodology, it is challenged by a greater need for cultural authenticity than for universal applicability. However, a comparison of culturally sensitive efforts can identify generally applicable principles, although just how those are enacted must again be meaningfully translated into the matrix of local living.

Kazarian and Evans concluded that although health psychologists have made important contributions to developing a more holistic approach to health, they have not effectively integrated culture into their approach. Their emphasis on the role of cultural health psychology in the creation of culturally competent health professionals is timely given the recent enthusiasm for a broader conceptualization of health, referred to as "global health." Global health is an attempt to address health problems that transcend national boundaries, may be influenced by circumstances and experiences in other countries, and are best addressed by cooperative actions and solutions. The underlying assumption is that the world's health problems are shared problems and, therefore, are best tackled by shared solutions. As Lee wrote, "Health care systems are traditionally structured, regulated, and delivered at the national level. Health research is correspondingly focused on the state, sub-state and inter-state levels. . . . There is a clear desire for a fuller understanding of the links between globalisation and health, beginning with the development of relevant conceptual frameworks and rigorous empirical analysis."

The health impact of globalization with regard to HIV/AIDS, coronary heart disease, and many other public health problems can be all the better appreciated by a comparative analysis—both conceptual and empirical—of how these difficulties affect people, and are affected by people, in a cultural context around the world. Beyond the needs for research, analysis, and policy development to which cross-cultural health psychology should contribute, there is also enthusiasm for new curricula that place global health center stage in the education of health professionals. With the increasing complexities of the contexts that may now be considered to constitute "cross-cultural" health psychology, and with greater recognition of the persuasive influence of global forces on human health, the value of cross-cultural health psychology may be increasingly realized in its contribution to addressing the challenges of global health.

See Also the Following Articles

Acculturation ■ Cross-Cultural Psychology, Overview ■ Health and Culture ■ Homosexuality

Further Reading

Aboud, F. E. (1998). *Health psychology in global perspective*. Thousand Oaks, CA: Sage.

Adler, N. E., & Ostrove, J. M. (1999). Socioeconomic status and health: What we know and what we don't. *Annals of the New York Academy of Sciences*, *896*, 3–15.

Airhihenbuwa, C. O. (1995). *Health and culture: Beyond the Western paradigm*. Thousand Oaks, CA: Sage.

Barlow, J., Wright, C., Sheasby, J., Turner, A., & Hainsworth, J. (2002). Self-management approaches for people with chronic conditions: A review. *Patient Education & Counseling*, *48*(2), 177–187.

Berry, J. W. (1997a). Cultural and ethnic factors in health. In A. Baum, S. Newman, J. Weinman, R. West, & C. McManus (Eds.), *Cambridge handbook of psychology, health, and medicine* (pp. 98–103). Cambridge, UK: Cambridge University Press.

Berry, J. W. (1997b). Immigration, acculturation, and adaptation. *Applied Psychology: An International Review*, *46*, 5–68.

Berry, J. W., Poortinga, Y. H., Segall, M. H., & Dasen, P. R. (2002). *Cross-cultural psychology*. Cambridge, UK: Cambridge University Press.

Blankenship, K. M., Bray, S. J., & Merson, M. H. (2000). Structural interventions in public health. *AIDS*, *14*(Suppl. 1), S11–S21.

Brown, L., Macintyre, K., & Trujillo, L. (2003). Interventions to reduce HIV/AIDS stigma: What have we learned? *AIDS Education & Prevention*, *15*(1), 49–69.

Campbell, C. A. (2002). Prostitution, AIDS, and preventive health behavior. *Social Science & Medicine*, *32*, 1367–1378.

Fogarty, L., Roter, D., Larson, S., Burke, J., Gillespie, J., & Levy, R. (2002). Patient adherence to HIV medication regimens: A review of published and abstract reports. *Patient Education & Counseling*, *46*(2), 93–108.

Glanz, K., Lewis, F. M., & Rimer, B. K. (Eds.). (1997). *Health behavior and health education* (2nd ed.) San Francisco: Jossey–Bass.

Green, L. W., & Kreuter, M. W. (1999). *Health promotion planning: An educational and ecological approach* (3rd ed.). Mountain View, CA: Mayfield.

Halpern, D., & Nazroo, J. (2000). The ethnic density effect: Results of a national community survey of England and Wales. *International Journal of Social Psychiatry*, *46*, 34–46.

Hawe, P., & Shiell, A. (2000). Social capital and health promotion: A review. *Social Science & Medicine*, *51*, 871–885.

Hermans, H. J. M., & Kempen, H. J. G. (1998). Moving cultures: The perilous problems of cultural dichotomies in a globalizing society. *American Psychologist*, *53*, 1111–1120.

Kazarian, S. S., & Evan, D. R. (Eds.). (2001). *Handbook of cultural health psychology*. San Diego: Academic Press.

Lee, K. (2003). *Health impacts of globalisation*. Basingstoke, UK: Palgrave Macmillan.

Lindau, S. T., Laumann, E. O., Levinson, W., & Waite, L. J. (2003). Synthesis of scientific disciplines in pursuit of health: The Interactive Biopsychosocial Model. *Perspectives in Biology & Medicine*, *46*(Suppl. 3), S74–S86.

MacLachlan, M. (Ed.). (2001). *Cultivating health: Cultural perspectives on promoting health*. Chichester, UK: Wiley.

MacLachlan, M. (2004a). *Culture and health: A critical perspective* (2nd ed.). Chichester: Wiley.

MacLachlan, M. (2004b). Culture, empowerment, and health. In M. Murray (Ed.), *Critical health psychology*. London: Sage.

MacLachlan, M., Smyth, C. A., Breen, F., & Madden, T. (in press). Acculturation and mental health in modern Ireland. *International Journal of Social Psychiatry*.

McDonald, H. P., Garg, A. X., & Haynes, R. B. (2002). Interventions to enhance patient adherence to medication prescriptions: Scientific review. *Journal of the American Medical Association*, *288*, 2868–2879.

Merson, M. H., Dayton, J. M., & O'Reilly, K. (2000). Effectiveness of HIV prevention interventions in developing countries. *AIDS*, *14*(Suppl. 2), S68–S84.

Mulatu, M. S., & Berry, J. W. (2001). Cultivating health through multiculturalism. In M. MacLachlan (Ed.), *Cultivating health: Cultural perspectives on promoting health* (pp. 15–35). Chichester, UK: Wiley.

National Research Council. (2001). *New horizons in health: An integrative approach*. Washington, DC: National Academy Press.

Nazroo, J. Y. (2003). The structuring of ethnic inequalities in health: Economic position, racial discrimination, and racism. *American Journal of Public Health*, *93*, 277–284.

Parker, R. G., Easton, D., & Klein, C. H. (2000). Structural barriers and facilitators in HIV prevention: A review of international research. *AIDS*, *14*(Suppl. 1), S22–S32.

Rivers, K., & Aggleton, P. (2000). HIV prevention in industrialized countries. In J. L. Peterson, & R. J. DiClemente (Eds.), *Handbook of HIV prevention* (pp. 245–263). New York: Kluwer Academic/Plenum.

Schneiderman, N., Antoni, M. H., Saab, P. G., & Ironson, G. (2001). Health psychology: Psychosocial and biobehavioral aspects of chronic disease management. *Annual Review of Psychology*, *52*, 555–580.

Swanson, B., Keithley, J. K., Zeller, J. M., & Cronin-Stubbs, D. (2000). Complementary and alternative therapies to manage HIV-related symptoms. *Journal of the Association of Nurses in AIDS Care*, *11*(5), 40–60.

Taylor, B. (2001). HIV, stigma, and health: Integration of theoretical concepts and the lived experiences of individuals. *Journal of Advanced Nursing*, *35*, 792–798.

Tuldra, A., & Wu, A. W. (2002). Interventions to improve adherence to antiretroviral therapy. *Journal of Acquired Immune Deficiency Syndromes: JAIDS*, *31*(Suppl. 3), S154–S157.

UN Program on HIV/AIDS. (1999). *Sexual behavioural change for HIV: Where have theories taken us?* Geneva, Switzerland: Author.

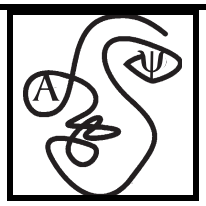
UN Program on HIV/AIDS. (2002). *Report on the global HIV/AIDS epidemic*. Geneva, Switzerland: Author.

Wilkinson, R. G., & Marmot, M. G. (Eds.). (1998). *Social determinants of health: The solid facts*. Geneva, Switzerland: World Health Organization.

Wilson, B. D., & Miller, R. L. (2003). Examining strategies for culturally grounded HIV prevention: A review. *AIDS Education & Prevention*, *15*(2), 184–202.

World Health Organization. (2002). *The World Health Report 2002: Reducing risks, promoting healthy life*. Geneva, Switzerland: Author.

Yudkin, J. S., Bayley, O., Elnour, S., Willott, C., & Miranda, J. J. (2003). Introducing medical students to global health issues: A bachelor of science degree in international health. *Lancet*, *362*, 822–824.



History of Applied Psychology, Overview

Helio Carpintero

Compultense University, Madrid, Spain

1. Introduction
2. Early Landmarks in Applied Research
3. The Institutionalization of the Field: Emerging Societies of Applied Psychology
4. Significant Problems in Early Applied Research
5. The Influence of World War II (1939–1945)
6. Theory and Practice: The Influence of the Era of Schools
7. Applied Psychology in the Second Half of the 20th Century
8. General Trends
   Further Reading

GLOSSARY

evolutionism A doctrine that maintains the formation of complex real entities from simpler ones, through an evolving natural process that brings in differentiation and heterogeneity in more homogeneous situations.

institutionalization The process of creating certain social structures that will be in charge of some definite roles originally solved by people's personal and original behavior.

intervention A concrete action that tries to introduce some changes in a given situation, usually planned and devised according to some previous theory, and adapted to the here-and-now peculiarities.

positivism A philosophical doctrine, first formulated by French A. Comte (19th century), that maintains the primacy of scientific knowledge and the need for building society upon a scientific worldview.

school A group of persons working on a certain field, according to some idiosyncratic views, and usually organized in a master–disciples structure.

theory A conceptual construction that tries to represent a certain group of phenomena in order to explain and predict them.

1. INTRODUCTION

Theoretical and applied psychology have frequently been considered as two completely different ways of approaching and conceiving of mental and/or behavioral phenomena. Their lines of development have met one another from time to time. But most frequently they have remained unrelated, as different questions, instruments, and methodologies have dominated in each of them, creating two different worlds.

Interest in psychological questions goes back to ancient times, as proven by philosophical reflections and literary creations. The diversity of characters and the description of typical reactions to a variety of situations form the core of countless literary masterpieces all over the world. Nonetheless, neither philosophical reflections nor literary fantasies enter our tableau. Psychology conceived as an empirical science is little more than a century-old enterprise. For didactical purposes, Wilhelm Wundt's laboratory at Leipzig

©2004 Elsevier Inc.
All rights reserved.

University (1879) is credited with the honor of representing the first event in its history.

This new science tried to reformulate old questions about the nature of the human mind and its creations—knowledge, volition, choice, consciousness—and its connection with the biological organism that owned such a mind. Two main ideas dominated the second half of the 19th century: evolution and science (the third one would be democracy). While evolutionary theory brought forth the idea of a continuity between humans and the animal kingdom, positivism stressed the value of scientific knowledge in all empirical domains and stimulated the creation of a natural science of the human mind. In what relates to the third idea, democratic values would soon demand the application of that new knowledge to solve the problems of societies developing under the impulse of the industrial revolution.

Two main lines, biological and sociological, characterize the early developments of psychological research. While certain groups focused on physical and physiological determinants of human behavior, others insisted on socio-historical factors mediating mental processes. On the one hand, work in laboratories and theory construction were stressed; on the other hand, researchers relied on social and ethnological comparisons and became largely interested in practical knowledge.

In the early days of the 20th century, applied research was not an easy task. In the United States, Edward B. Titchener (1867–1927), a "pure scientist," strongly opposed psychologists emerging from laboratories and facing daily life problems. He feared that their theories were not yet prepared for such a huge task. But demands could not wait unattended for long. Moreover, funds for research soon became dependent on practical results accredited by scientists. Little by little, applications began to follow theory, and science gave way to practical interventions; scientists were accompanied by practitioners, and psychology became a science as well as a profession, as it is at present.

Since its beginnings, interventions in psychology were characterized by the need for combining both a scientific background and information on the idiosyncratic characteristics of every practical problem. Applied science, in the case of psychology, has not only consisted of translating general principles to concrete circumstances, but also of creating new theoretical models that are fully in accordance with concrete data, trying to meet the particular demands, and gaining through such intervention new theoretical insights. Such circular interactions between research and theory were clearly stated by Kurt Lewin in his "action research" doctrine and have been frequently found by researchers in most of the practical fields.

Developments in the field have been deeply influenced by national trends. Progress cannot be reduced here to a single common process covering all countries and specialties. Three main areas did emerge, however, as their problems were found in all well-developed societies: (1) children's adaptation to school, (2) worker's adaptation to job requirements, and (3) people's adaptation to the requirements of "normal" life in a social world. These areas led to the formation of school and educational psychology, industrial and organizational psychology, and clinical psychology, respectively. As time passed, the mapping of specialties became more and more complicated.

Certainly, among all these differences there are some common traits in these applications that may serve to unit and define the field of applied psychology. These traits will be considered after a general discussion of the field.

2. EARLY LANDMARKS IN APPLIED RESEARCH

While the establishment of Wundt's laboratory is a landmark that represents the first step, there is no formal beginning for applied psychology as such. In each specialty, however, the first steps have had some unique characteristics of their own that will be considered here.

2.1. Clinical Psychology

It is not easy to disentangle clinical psychology from the fields of mental hygiene and psychiatry, but it might be accepted that problems and deficits in mental aptitudes can be seen as psychological problems, whereas deep mental disorders require medical treatment. The former will impair normal adaptation to life, while the latter will prevent normal functioning of the person among his relatives and in his private world.

The first steps in the field of clinical psychology stem from the introduction of the idea that a mental process and its abnormal functioning could be treated through behavioral/mental means, instead of using somatic or chemical procedures. An empirical approach to the study of mental disorders was demanded in this new science of the mind, as these disorders began to be seen as natural disturbances capable of being treated. Some epochal events took place in the year 1896. In Germany, a student of Wundt, psychiatrist

Emil Kraepelin (1856–1926), first proposed the introduction of the psychological laboratory and the experimental study in psychiatry, in a seminal paper that has been called the first German program for clinical psychology. Kraepelin divided illnesses into two categories: those due to endogenous factors and those caused by exogenous ones; the former were viewed as usually incurable, while the latter had a better prognosis. Based on empirical grounds, he also worked on a classification of mental disorders that served for decades as a nosological system.

Also in 1896, an Italian psychiatrist, Giulio Cesare Ferrari (1868–1932), with G. Guicciardi, published a paper on mental tests that in their opinion could be employed in psychiatry to improve diagnosis. They tried to continue the pioneer work of Gabriele Buccola (1854–1885), who, working with patients suffering from mental disorders, found interesting variations in their reaction time measures. Ferrari and Guicciardi examined attention, associative abilities, and other mental capacities in patients in order to more accurately characterize their pathologies.

And it was in the same year when, in the United States, Lightner Witmer (1867–1956), another student of Wundt and Cattell, founded the first psychological clinic, at the University of Pennsylvania. Witmer became an apostle of the psychological intervention. At the American Psychological Association (APA) convention of that same year (1896), he maintained that psychologists had to assume two tasks: to study mental processes and to apply their knowledge to solve personal problems and social needs. He had to face the negative attitude of his colleagues, who did not wish to accept practical responsibilities. He decided to organize regular training courses at his university for people interested in applications of psychology, with case presentations and formation in diagnostic techniques, and also founded a journal, *The Psychological Clinic* (1907), that claimed to promote "orthogenics," the knowledge of the correct development of individuals, both in physical and mental capabilities.

A different approach was carried out by William Healy (1869–1963), a physician who treated juvenile delinquency as a phenomenon he believed was rooted in certain psychopathological factors. He created a battery of proofs and applied them to young offenders, trying to determine their personalities and social factors in order to implement preventive and therapeutic interventions. His first clinic was established in Chicago (1909) and then he moved to Boston, where he retired.

In sum, demands for help from psychologists, coming from the social world, were based on the assumption that the new science of psychology, which was dealing with the normal functioning of the mind, could also say something about those pathologies now interpreted as mental disorders.

2.2. Educational Psychology

The 19th century saw the democratic rise of masses in Western countries, and basic education and training spread out to all individuals and classes. Many people became teachers, but they soon discovered that their job was not an easy one. Individuals are endowed with unequal learning capacities, and those prevented from normal schooling caused great disturbances in the classroom. Because their need for special treatment was ignored by most teachers, they turned to psychologists for orientation and help. Whereas old-fashioned schooling was based on memory and rote learning, some new tendencies stressed the role of an active process of learning and understanding, a true psychological topic. Teachers viewed psychology as the scientific basis for their own art of teaching.

The association paradigm became largely influential. In 19th century Germany, Johann F. Herbart's (1776–1841) association became the basis of the new science of education; in the United Kingdom, interest in these matters went back to empiricist philosophers, from John Locke to James Mill, whose ideas inspired 19th century psychologists such as Alexander Bain (1818–1903) and James Sully (1842–1923). Sully made careful empirical studies of language, perception, memory, and other faculties in children, and wrote one of the first textbooks, *Teacher's Handbook of Psychology* (1886).

French pedagogue Gabriel Compayré (1843–1913) wrote another textbook for teachers that included not only theoretical concepts on association and other mental aspects, but also practical topics, such as children's plays, how to improve curiosity and memory, and the technology of teaching. His book was translated into Spanish by the great Argentinian educator and president, D. F. Sarmiento, and became influential among Latin American teachers and pedagogues.

Largely under the influence of Wundt's psychology, an experimental pedagogy began to spread in Germany. Scientific knowledge of child development, based on observation and experiments, was systematized in the pioneering work of Wilhelm Preyer (1841–1897),

The Soul of a Child. Pedagogues such as Wilhelm A. Lay (1862–1926) and Ernst Meumann (1862–1915) multiplied their efforts to provide psychological support to educational practices. Some of them stressed the need for building a global knowledge of child life, and a first draft was accomplished by German Oscar Chrismann in 1889, with his proposal of a "paidology"—or "pedology"—conceived of as a general science of the mental and physical development of the child. Centers for research and teaching were then founded, such as the Institute for Pedagogy and Experimental Psychology of Leipzig, and the Institute for Applied Psychology, created by William Stern and Otto Lipmann in Breslau, both founded in 1906. Specialized journals such as *Zeitschrift fuer Kinderforschung* (1895) also began to publish the results of their research.

Some educators spoke up against undue expectations of psychology. William James (1842–1910), in *Talks to Teachers* (1899) made clear the distinction between the science of psychology and the art of teaching, affirming that "sciences never generate arts directly out of themselves." But some of his students, like Granville Stanley Hall (1846–1924), assembled a large body of knowledge about lifespan psychology (e.g., childhood, adolescence, senescence), and wrote a seminal work on "The Content of Children's Mind on Entering School" (1883) in which he empirically explored the average knowledge of children in his society. This field soon became an active area of research. Hall expanded that paper into a book (1894), and founded the journal *Pedagogical Seminary* (1891), later renamed the *Journal of Genetic Psychology*.

Interactions between psychology and education multiplied rapidly, and benefited from the growing knowledge about children promoted by evolutionism. In Belgium, paidological topics received great attention, especially topics related to mental retardation. A. Sluys, J. Demoor, and Ovide Decroly (1871–1932), among others, made lasting contributions that became influential all over Europe. The latter encouraged the adoption of the New School model, which promoted the ideals of active learning and the view of the school as a laboratory.

Great attention was also paid to these topics in socialist countries. In Russia (in those days the Soviet Union), many centers oriented toward experimental study of child development and school education were founded, as the building of a "new man" was a political goal according to Marxist views. Lev Vygotsky (1896–1934), Pavel Blonsky (1884–1941), Aron Zalkind (1888–1936) and others worked on this, stressing the social and practical roots of human consciousness. A Congress of Pedology gathered in 1928, a journal (*Pedology*, 1928–1932) was founded, and activity grew rapidly, until July 4, 1936, when a decree from the Communist Party headed by Josef Stalin halted testing and deemed nearly all psychological practices counterrevolutionary activities.

The ideals on active education were at the core of an institution created in Geneva, Switzerland in 1912 by Edouard Claparède (1873–1940), Pierre Bovet, and some collaborators, the Institute Jean Jacques Rousseau, which soon became a model for all who were interested in those topics. Claparède offered an overview of the field in his *Psychologie de l'Enfant et Pédagogie Expérimentale* (Child Psychology and Experimental Pedagogy, 1905), in which he stressed the idiosyncratic nature of child life and the need of an educative system that is well adapted to the individual nature of every child and capable of giving effective guidance to students.

Some early studies had great impact on the field, such as research on formal discipline and its effects on concrete learning. In the United States, Edward L. Thorndike (1874–1949) and Robert S. Woodworth (1869–1962) performed experiments on the transfer of training that showed poor or negative aftereffects of a first learning task upon another following it. In doing so, they destroyed the belief that some benefits might be obtained from disciplines such as Latin or technical drawing upon other matters included in the curriculum. But the study of individual differences among children and a more detailed knowledge of the evolution of intellectual abilities and interests enabled teachers to adjust their teaching to the children's capabilities.

Apart from these descriptive studies, more effective instruments for assessing children in schools were demanded. In response, some psychologists built instruments for screening deficits and assessing effective teaching. They initiated the application of proofs and measures to education. This will be discussed in a following section on testing.

2.3. Early Work in Organizational Settings

Western societies have created all sorts of organizations in which the principles of impersonal work, labor division, cooperation, and leadership applied. In modern states, people operate under the leadership of certain elites, and technology introduced during the industrial revolution is changing the world through the work of

people working together. In the early decades of the 19th century, the need for a science of social movements and phenomena stimulated the creation of sociology, an accomplishment of French theorist Auguste Comte (1798–1857), the founder of positivism. Soon after, theorists tried to explain some dimensions of social life as mental phenomena, as well explain some psychological facts by social forces. Interactions between these fields multiplied, and psychological knowledge of individuals operating in a group or organization seemed more and more convenient. Aspects such as job–person fit, efficacy, performance quality, accident prone operators, and social competition were not without economic consequences. Enterprises in search of good operators soon incorporated psychological testing devices in order to find reliable and valuable workers. Psychology appeared as a science dealing with problems closely related to person–job adaptation, and soon society asked psychologists for help in the task of improving industries and social organizations.

The concept of applied psychology was first delineated by German William Stern (1871–1938). According to him, it was "the science of psychological facts that are relevant by their practical utility," including both assessment and intervention (Dorsch, 1963, p. 9). Early essays on this topic are characterized by their spontaneity and diversity; frequently they were written by laymen interested in practical questions. For instance, Italian physiologist Angelo Mosso (1846–1910) stressed the importance of analyzing the fatigue curves of workers to improve their performances; his findings and methods spread rapidly in Europe. On the other hand, a U.S. engineer, Frederic W. Taylor (1856–1915), working for important steel companies, studied their production from the point of view of the productive behaviors and their timing and morphologies, giving place to a movement called the "scientific organization of labor," that applied the scientific methodology to the study of people's efficiency at work, and the ways to improve it. Although not a psychologist himself, he applied psychological techniques—reaction time, movement studies, cooperation among workers, economic incentives—in trying to combine personal satisfaction with greatest worker efficiency at the risk of being judged an inhumane manager. His proposals introduced a new concept to industries, and while social movements and trade unionism shook developed societies, the problems of people at work began to be viewed under the light of psychology. Taylorism paved the way for a more specialized organizational psychology.

The new climate brought new fruits. Frank Parsons created the Boston Vocation Bureau (1908), convinced that a test for discovering motives and aptitudes in individuals could improve their professional adaptation. As a philanthropist, he wanted to help people in search of a job, and his center considered many aspects: people's abilities, technical and moral training needed, self-understanding, and the material conditions of the job.

A world becoming dominated by technical apparatuses and procedures demanded efficient people for using them. Car driving and traffic security also stimulated the screening of potential accident prone drivers via psychological testing, a task in which German researchers Walter Moede (1888–1958) and Curt Piorkowski (1888–1939) did significant work employing driving simulators. Motion studies of bricklayers done by Lillian M. Gilbreth (1878–1972) and Frank Gilbreth (1869–1924), and of typist's abilities done by Jean M. Lahy (1872–1943), among others, stimulated the research, but it was Hugo Münsterberg (1863–1916), the "father of psychotechnics," who was the first to face the problem in its entirety and gave a certain solution.

Trained under Wundt, then working with James at Harvard University, Münsterberg conceived of psychotechnology as the "science of practical applications of psychology in the service of culture" (1914). His books on industrial efficiency, teaching, and eyewitness testimony offered a broad view of psychology as a practical science. He devised proofs and testing situations that proved useful for personnel selection, and received much support from businessmen and industry managers. His works describe the new field of psychological intervention as a coherent whole, in which demands were idiosyncratic, but techniques were interspecific and largely versatile in themselves. Person, situation, and means-ends relations were the significant elements to take care of in every practical intervention.

2.4. Early Testing

Early psychophysics—the scientific study of stimuli and the correlated sensory experience—found systematic variations in sensory and motor abilities among individuals. One of the pillars of Darwin's evolutionism was the existence of spontaneous variations among individuals who were unequally endowed, in order to cope with the challenges of their environment and subject to natural selection.

Francis Galton (1822–1911), who was deeply influenced by evolutionism, studied such variations and devised a battery of tests to measure differences in various types of capacities among subjects. He applied such proofs in an anthropometric laboratory at the International Health Exhibition of South Kensington in 1884, and created some statistics to sum up his data.

James McKeen Cattell (1860–1944), a U.S. student of both Wundt and Galton, had a long-lasting invention: the construction of standard proofs to compare the performance of different people. In 1890, he coined the expression "mental test" for measuring mental capacities, which spread out all over the world and became a symbol for psychology as a profession. His presidential address at the APA convention in 1895 enforced the need to present to society not only scientific but also material successes in the practical field, in order to enhance the status of psychology; the simplest experimental situation in which a person could be studied and measured was the test. For that reason, tests grew without limitation in the 1920s and 1930s.

In Germany, Hermann Ebbinghaus (1850–1909) devised a completion test, in which some missing letters had to be found, for assessing the effects of fatigue in school children, a type of proof frequently used since then. But it was Alfred Binet (1857–1911) who, with his collaborator Victor Henri (1872–1940), made an epochal contribution: the metric scale for intelligence (1905), a worldwide test for assessing intellectual abilities, allowing the determination of the relative position of a child (his "mental age") among a certain population. Moreover, intelligence soon appeared to be highly correlated with most mental abilities, as a kind of broad cognitive strength or general factor influencing all types of performances in which those abilities operate, as found by Charles Spearman (1863–1945).

Evaluation and comparisons between individuals became then possible. Age was the basis for comparisons in Binet's system; a further step to get a stable index across ages came from German psychologist William Stern (1871–1938), who introduced the concept of the intelligence quotient (IQ), the ratio of mental age to chronological age, a score that seemed to remain unchanged along the years. Its measure became a highly valuable index of a person's capabilities and a good predictor of future performances.

Benefits from the possibility of measuring intelligence began to appear everywhere. School was only one field interested in such measures; social life, job selection, even clinical practice soon could profit in different ways from this measurement. The various branches of psychological practice began, little by little, to cross-fertilize each other, and their methods and purposes approached each other, producing a certain unification of the applied field as a whole.

2.5. Interactions across Specialties

Three main facts caused the explosion of applications of psychology to social problems: the usefulness of intelligence testing for successful prediction of future performance; the growing need of a scientific explanation for the behavior of people in organizations shaken by social and political unrest; and the rise of a clinical view of abnormality in the light of psychological conflicts of personality, as in the contributions of Sigmund Freud (1856–1938). All these historical innovations converged in viewing mind and behavior as the Gordian knot of the problems of modern societies. Psychologists were asked to respond to those problems, and an external event, World War I (1914–1918), provided the opportunity to prove their social efficiency. As Cattell wrote, it "put psychology on the table," and nobody could ignore it.

War and military needs can be considered as some of the strongest stimuli for developing psychology in societies. World War I gave a great impetus to all the applied branches; the areas of assessment and personnel selection, and also clinical intervention upon war-caused psychopathologies, quickly developed. Three main problems emerged during wartime: war neuroses (then called shell-shock), the health of munitions workers (menaced by industrial fatigue), and military personnel selection.

In the United States, psychologists from the APA, headed by R. Yerkes (1917) and considered as a part of the medical division, worked out a test battery (the Army Alpha for literates, and Army Beta for illiterates) for assessing soldiers. Both proved to be effective. Examinations were given to more than one and a half million people, and since then, tests have been an indispensable part of any personnel selection. The war also was the motivation for the creation of some clinical instruments, like the personal data questionnaire (1918) devised by R.S. Woodworth (1869–1962) to detect symptoms of a neurotic tendency. With the end of the war, these particular activities greatly diminished.

The war also promoted the introduction of psychology as part of the Army services in Germany. There some regular positions for psychologists were created

in 1925, mainly oriented toward officer selection. This number increased after 1935, as military psychology became synonymous with psychology under the Nazi government. In Great Britain, the war brought the problem of industrial fatigue to the forefront. In munitions factories, people had doubled their working time, and physical and mental disorders rapidly appeared, affecting the subjects' health and efficiency. To face the situation, a Health of Munitions Workers Committee was set up by the government in 1915; it was renamed the Industrial Fatigue Research Board (1917) and then became the National Institute of Industrial Psychology (1921), headed by Charles S. Myers (1873–1946). Similar efforts were carried out in other countries. Airplane pilots were urgently needed, and they were examined by psychologists. In Italy, Father Agostino Gemelli (1878–1959), founder of the first psychophysiological laboratory of the Italian Army, did his best to develop applied psychology in his country in the 1920s and 1930s; in France, Camus and Nepper (1915) did similar work.

Around the same time, the first centers for applied psychology began to appear. In the United States, a psychology department was established in 1915 at the Carnegie Institute of Technology, headed by Walter Van Dyke Bingham (1880–1952), and a year later, Walter Dill Scott (1869–1955) was appointed as professor of applied psychology, the first to receive that title in that country. Vocational counseling, industrial efficiency, and related topics began to be treated as academic matters, so the time was ripe for institutions to specialize in those questions. For example, the Kaiser Wilhelm Institut, the Laboratory for Industrial Psychotechnology from Charlottenburg, the Institute for Work in Moscow, the Institute of Professional Orientation of Barcelona, the Institute Jean-Jacques Rousseau at Geneve, the Psychological Laboratory at the Catholic University of Sacred Heart (Milan), the Institute of Scientific Organization at Warsaw, the Intercommunal Center for Professional Guidance at Bruxelles, the Laboratory for Ergonomics of the Institute of Higher Studies of Belgium, the Laboratory of Experimental Psychology of the School for Higher Studies at Paris, the Center for Professional Orientation, and the Bos Institute (the Netherlands), among others, became active centers for training, testing, and research. These institutions and centers were created to face social demands, and utility was the dominant criterion accepted by researchers, although in the long run many theoretical contributions were produced from such practical work.

3. THE INSTITUTIONALIZATION OF THE FIELD: EMERGING SOCIETIES OF APPLIED PSYCHOLOGY

The period that followed World War I brought with it the normal problems in a rapidly growing field without any internal regulation, mainly under external pressures. Psychology was acknowledged as a useful instrument for social work, and its techniques for personnel selection spread both to private firms and to governmental agencies.

To face the new demands, psychologists tried to set up several professional networks. The need for coordination and the convenience of employing common instruments for assessment stimulated the foundation of an international society for applied psychology, was the *Association Internationale des Conferences de Psychotechnique* (International Association for Psychotechnology). Rooted in Europe but open to the whole world, it held its first meeting, chaired by Swiss Edouard Claparède (1873–1941) in Geneva in 1920; in 1955 it changed its name to the International Association of Applied Psychology (IAAP). The society provided a useful platform for cooperation, but was greatly affected by the political circumstances that led to World War II; 15 years passed between the 8th Congress, held in Prague in 1934 and the 9th, held in Bern in 1949. Since then, it has gathered regularly. The IAAP has been headed by leading scientists: Edouard Claparède (1920–1941), Henri Piéron (1947–1953), Clifford Frisby (1953–1958), Morris S. Viteles (1958–1968), Gunnar Westerlund (1968–1974), Edwin A. Fleishman (1974–1982), Claude Lévy-Leboyer (1982–1990), Harry C. Triandis (1990–1994), Bernhard Wilpert (1994–1998), Charles D. Spielberger (1998–2002), and Michael Frese (2002–present).

Many other international associations have been created to serve the multiple interests of psychologists. The International Union of Psychological Science (IUPsyS) was created in 1951, grouping together many previous national associations. While its roots go far back (a first International Congress of Psychology took place in Paris in1889), the IUPsyS now represents psychology in international councils of science and culture. Also, the Interamerican Psychological Society (*Sociedad Interamericana de Psicologia*, SIP), was founded in 1951 to cover the needs of psychologists from the entire American continent; it is very active at present. Many other specialized societies have flourished in parallel to the growth of the profession and the scientific field.

An outstanding society as measured by its size and activities is the American Psychological Association (APA). Founded by G. S. Hall in 1892, in its beginnings it covered only scientific matters. Therefore, applied oriented psychologists created the American Association of Clinical Psychologists in 1917, but finally this group was included in the APA in 1919 as a clinical section dedicated to professional issues. A new group called the American Association for Applied Psychology was founded in 1937 to better serve practical interests. It edited a journal (*Journal of Consulting Psychology),* but eventually it too joined the APA in 1945, forming a strong society, organized in various divisions, oriented to both science and profession.

The growing complexity of the applied field necessitated a communication network that could disseminate specialized information. During the years of war, some periodicals began to appear. In Germany in 1916, the Institute for Applied Psychology of Leipzig edited the *Zeitschrift fuer Angewandte Psychologie* [Journal of Applied Psychology] under the care of William Stern and Otto Lipmann. In 1917 in the United States, G. Stanley Hall initiated the publication of the *Journal of Applied Psychology*, another one of the leading journals in the field. It was soon managed by James P. Porter (1920–1942), and later owned by the American Psychological Association. Other significant journals have been the *Revue de Psychologie Appliquée* [Review of Applied Psychology] (1920), and *Human Factor* (1932), under the sponsorship of the National Institute of Industrial Psychology and the initial editorship of C. S. Myers. Since then, the network has developed enormously, and the creation of the *Annual Review of Psychology* (1950) and more recently the data banks that have been built in many specialties have introduced a new order into the information chaos.

4. SIGNIFICANT PROBLEMS IN EARLY APPLIED RESEARCH

Some basic problems arose during the evolution of applied psychology. The following sections briefly describe some examples of these problems.

4.1. Inborn versus Learned (Acquired) Abilities

Theorists had to decide whether human abilities were innate or learned; testing procedures and measurements raised suspicions and opposition among socially involved groups on political grounds. Large differences in measurements between various classes and ethnicities raised political doubts about the fairness of procedures that, in fact, introduced discrimination between races, classes, and individuals. In addition, an increasing number of dimensions of life were subject to study under this methodology.

One subject of intense debate was intelligence. Did the result of testing stem from the inborn characteristic of the person or from the social conditions of the person's life?

Early on, Francis Galton (1822–1911) was inclined toward a genetic solution to the dilemma. In the United States, H. H. Goddard (1866–1957), a Binet follower, reinforced that opinion. He compared two lines of descendants, both legitimate and illegitimate, flowing from a common ancestor and examined their achievements and life records. He concluded that hereditary factors explained most of the variance, as the legitimate family was largely superior to the other one in intellectual and moral qualities.

Different races and cultures were also compared on the basis of their performance on intelligence tests. Data from the Army Alpha and Beta tests were studied and reanalyzed in the United States. Significant differences in IQ among people from various ethnic origins were supposedly found. Among other questioned results, the intellectual weakness of African Americans seemed to be firmly established on these grounds. According to Garrett (1951), "Negro soldiers scored lower than whites on Alpha, southern Negroes lower than northern Negroes." Following these results, democracy was criticized by some groups on the grounds of offering political equality to unequally mentally endowed people. Racist arguments seemed to flow from psychological data. As a consequence, U.S. federal law regulating immigration excluded all "persons of constitutional psychopathic inferiority."

Other voices joined the choir. In Europe, data obtained by H. English and others suggested a strong correlation between economic level and IQ and pointed to non-inborn factors in intelligence scores. The discussion was deeply affected by political prejudices and attitudes. IQ testing was then criticized on political grounds. In the United States, an influential journalist, W. Lippmann (1899–1974), strongly criticized the testing procedures for taking scores as true measures instead of rough estimates of variable qualities in subjects.

In Russia, then under the Soviet regime, a ban on all testing activities in schools was imposed by the

Communist Party's Central Committee. They deemed such activities a "pedological perversion," considering tests as reactionary and anti-egalitarian techniques. Given the thesis of the social determination of consciousness caused by socioeconomic factors, the idea of inborn abilities was seen as unacceptable. Such a decision greatly impaired the development of empirical psychology in the country for decades.

As is well known, this debate has endured for an entire century and at certain moments has become intertwined with political attitudes. The works of A. R. Jensen (1967), H. J. Eysenck (1971), L. Kamin (1974), S. J. Gould (1981), R. Herrnstein, and C. Murray (1994) represent only a small part of the vast literature generated by this controversy.

Today, heredity and environment are placed in a completely different relationship, as genetic engineering has resulted in a complete mapping of genes. Proteins related to organic and mental functioning are now beginning to be discovered, and intervention at such microscopic levels will be possible in the future, opening the way to biological techniques of medical and psychological value.

4.2. Types of Measurement and Types of Diagnosis

A long-ranging debate commenced between those inclined to build diagnoses upon a collection of measures, including personality traits and abilities, and those who limited themselves to the study of the subjects' aptitudes directly related to the job they were applying for. In the early days, theorists such as German O. Lippmann (1880–1933) tended toward a holistic approach, while many others (e.g., Belgian A. Christiaens) favored the opposite view.

The debate evolved into the opposition between qualitative and quantitative (psychometric) procedures for assessing people. The former was frequently called clinical methodology, because it focused on concrete phenomena affecting a single subject (idiographic methods), while the latter was frequently referred to as statistical methodology, as far as it employed quantitative measures obtained from groups of subjects through psychometric tests (nomothetic methods). For instance, some theorists, such as Swiss cognitivist Jean Piaget (1896–1980), firmly rejected the use of psychometric proofs, convinced they gave wrong information about the child's mind, and systematically employed clinical procedures. Researchers also developed instruments to generate and appreciate qualitative experiences that might be directly related to mental forces and impulses—the projective techniques, which included well-known proofs such as Rorschach's inkblot test (1921) and Murray's Thematic Apperception Test (1938). Hermann Rorschach (1884–1922), a Swiss psychiatrist, wrote *Psychodiagnostik* based on the experimental study of characteristic answers to some inkblots provided by different criteria groups of psychiatric patients. On the other hand, many psychologists tried to devise laboratory situations closely related to those found in real situations.

Opposition to mental measurement also grew from theoretical grounds. Behaviorism, which largely dominated U.S. psychology between the 1920s and 1960s, banned all mental-based concepts from its system, especially from all its clinical topics. It stated that psychology should deal only with behavioral facts and laws, and each individual should be tested in definite settings, in order to establish those precise S-R associations determining his or her behavior. Adaptive or maladaptive habits were substituted for old personality traits, and behavioral changes were evaluated in order to assess the effectiveness of various procedures of behavior modification. New instruments based on observational procedures were developed, employing sampling recording of target behaviors in single subject ($n = 1$) cases; these techniques substituted for traditional tests and questionnaires.

4.3. Studying People or Analyzing Situations

Early psychotechnology tried to study the adaptation of a person to a situation by comparing the person's qualities to the requirements of each type of situation. Typologies and professions were its main fruits. But some unfitting data opened the door to the study of work as a global social situation. Surprising data emerged from some studies on work efficiency carried out by George Elton Mayo (1880–1949) and collaborators on employees of the Western Electric Company in Hawthorne, Chicago. An in-depth study with those workers showed that their efficiency at work was mainly dependent not on physical factors at the workplace but on social characteristics of the group—informal rules, leadership, social climate, expectations, and fears—and also from the fact that they were receiving attention from researchers, a fact that influenced the feelings of those

workers. Mayo, F. G. Roethlisberger, W. J. Dixon, and G. Homans placed emphasis on human relations in industry and stressed the relevance of the "human meaning" of work and the necessity of a holistic approach to work problems. Significant developments in this theory were obtained when extra-organizational variables, such as technological and market changes, religion, race, and social class, were also taken into account by the Chicago school (W. L. Warner and colleagues) and other groups.

All these experiences showed the need to study the characteristics and behaviors of groups. Gestalt psychologist Kurt Lewin (1890–1947) emerged as a leader in that field. His ideas influenced not only the general field of social psychology, but also the study of applied aspects, such as the relationship between school climate and teaching efficiency. Lewin, R. Lippitt, and R. K. White evaluated the effect of different group atmospheres—democratic, authoritarian, and *laissez faire*—upon school teaching and group interactions. Leadership proved to be an important variable, and various models of it have been devised to explain empirical data. Results from industrial settings were transferred to educational and clinical fields. Group psychotherapies have been applied since the 1920s by Alfred Adler (1870–1937) and coworkers, and group interactions proved fruitful in handling feeling and personality disorders. Well into the 1950s, group psychology developed a variety of forms and applications: for instance, T (training) groups were created that offered training in skills and interpersonal communication; encounter groups were employed for therapeutic discussions in humanistic psychotherapies. A special place in this line of thought belongs to the Tavistock Institute (UK) founded in 1947. It focused on the interactions between technological and psychosocial aspects of organizations, and tried to combine psychoanalytic concepts and the group dynamics as developed by the Lewinian group of theorists. A. K. Rice, E. L. Trist, W. R. Bion, and many others helped the development of this movement, and its journal, *Human Relations* (1947) maintained a comprehensive perspective in the field.

4.4. Intervention versus Goal-Oriented Research

As applied psychology rapidly developed, more and more concrete questions absorbed the interest of researchers, who began to specialize in a small range of topics and to sight of the entire field. An increasing distance began to appear between theory and applications, and as a result, some tried to restore the communication between both fields.

After World War II, clinical psychologists in the United States felt the necessity of combining practice with research. In a conference at Boulder, Colorado in 1949 they devised the so-called Boulder model, which recommended "the training of the clinical psychologist for research and practice, with equal emphasis on both" (Raimy, 1950). This model has been widely accepted, and is still in force among specialists.

Similarly, social psychologists doing research and intervention in social settings were warned by K. Lewin in 1948 that their interventions had to be organized "in a spiral of steps each of which . . . composed of a circle of planning, action, and fact-finding about the result of the action." In combining planning, action, and evaluation, researchers will adjust themselves to a Lewinian "action research" model that puts together theory and action into a single operation, a model assumed by most social scientists. These innovations perceived practice and theory as two dimensions of the same impulse toward knowledge.

5. THE INFLUENCE OF WORLD WAR II (1939–1945)

World War II deeply changed the social scene, and all aspects of human life, including psychology, felt its impact. Numerous institutions, researchers, and schools of thought were destroyed or silenced by the war. In Europe, an enormous number of scientists and intellectuals belonging to Jewish families were persecuted and even killed in countries placed under the Nazi regime. Psychologists were no exception. Many of them fled to the United States or to other countries, and European psychology, until then considered the best in the field, lost its position of leadership to the United States.

In Germany, some psychologists carried out research for the new regime. The German army established a *Wehrmacht* psychological division, which performed the selection of specialists and officers; the Reich Institute for Psychological Research and Psychotherapy trained "doctor psychologists" for psychotherapeutic labor. Psychologists found new professional perspectives not limited to the academic field; at

the same time, people teaching in universities had to show their loyalty to the racist ideology then dominating the country.

Likewise, many psychologists in the United States worked for the U.S. army, doing personnel selection, training, and psychotherapy. Limited but important research was also carried out on topics like morals, propaganda, and psychological warfare. Well-known researchers such as R. M. Yerkes (1876–1956) and E. G. Boring (1886–1968), among others, made great contributions to that effort; more than a thousand professionals worked full time to face the emergency. During that time, a new test (the Army General Classification Test) was administered to some 12 million soldiers. Its results favored an in-depth study on aptitudes and abilities.

The war also brought to the fore new problems and research topics for psychologists. Studies on visual perception in flight situations fostered new research in ecological theory like that carried out by James J. Gibson (1904–1979); new techniques for radar detection systems paved the way for a new psychophysics, based on signal detection and decision theory; vigilance and task monotony induced new research on stimulus deprivation and attention processing; race prejudice and antisemitism stimulated research on authoritarian personality by Theodor Adorno (1903–1969) and colleagues; and information theory, developed by Claude Shannon (1916–2001), and cybernetics, by Norbert Wiener (1894–1964), were theoretical advances born from practical needs that soon induced new developments in psychological theory as well as in the world of automation; post-traumatic stress disorders and the demand for help from clinical psychologists stimulated research in this field emphasizing the relationship between organic (somatic) reactions and social experiences. War tied together theory and practice, after all.

6. THEORY AND PRACTICE: THE INFLUENCE OF THE ERA OF SCHOOLS

Theoretical psychology experienced in the first half of the 20th century an extraordinary diversification of its viewpoints. That period has been frequently called the era of schools. Some viewpoints became largely dominant, and conflicts between them arose both in theoretical and practical questions. Behaviorism, Gestalt, psychoanalysis, and the humanistic approach were among the most influential perspectives in the United States; in the European countries, on the other hand, functionalism and factorialism became very influential. In certain cases, a special doctrine was dominant—as is the case of the Marxist-based Soviet psychology that monopolized Russia under the Soviet regime, or the Piagetian school that influenced European child and school psychology in the middle of the century. These schools of thought are briefly described in the following sections.

1. Behaviorism, created by John B. Watson (1878–1958), made behavior the proper object of study for psychologists, while excluding mind and consciousness as "subjective non-public entities" and, as such, unsuitable for science. Built upon an S-R (stimulus-response) paradigm, theory was conceived of as a guide for exerting technological control over behavior. Since its early days, behaviorism advocated significant applied research. Watson himself wrote a book on child training that was largely influential among U.S. families in the 1930s; he was also for many years a practitioner of commercial and advertising techniques.

Predicting and changing behavior became an essential function of psychologists; the practical value of such control gave birth to an entire field, "behavior modification," that applied the laws of learning and motivation to changing people's behavior. Pioneering work was carried out by one of Watson's students, Mary Cover Jones, who, based on Watson's ideas of conditioning, devised a successful treatment of a phobia in a child.

Behavior modification appears to be a case of a multiple discoveries that took place in the early 1950s. Burrus F. Skinner (1904–1990) in the United States, Joseph Wolpe (1915–1997) in South Africa, and Hans J. Eysenck (1916–1997) in Great Britain all had the same idea of conceiving mental disorders as learned (maladaptive) behaviors, subject to change for the better (re-learning) under adequate manipulation of effective variables. The new approach rapidly spread to the Western countries. In the United States, a large group of researchers oriented themselves toward the applied behavior analysis, devising interventions on social problems according to its principles. In Latin America, where psychoanalysis largely dominated in the 1940s and 1950s, behavioral techniques appeared as a greatly desired alternative in clinical practice. Under the influence of some U.S. specialists, the movement began to consolidate in the1960s in Brazil, Mexico, Chile,

Venezuela, and Colombia, and the Latin-American Association for the Behavior Analysis and Modification (ALAMOC) was founded in 1975. A similar phenomenon took place in European countries, where these techniques proliferated in the 1960s, and the European Association for Behavior Therapy was created in 1970. Adepts to these principles were oriented toward practice or toward theoretical questions. Techniques for handling stimuli and motives have been devised and applied to nearly every human situation—schools, family, industries, therapies, and so on. While very effective in multiple settings, behaviorism has also been criticized for its appearance of merely dealing with symptoms, with total neglect for the causes of behavior. After the emergence of the cognitive paradigm, a combined approach—the cognitive behavior modification—was generated; it focused on cognitive control and influence upon behavior as a means of rebuilding a healthy state.

2. Psychoanalysis, created by Austrian psychiatrist Sigmund Freud (1856–1939), has turned into one of the biggest intellectual and social constructions of the 20th century, influencing all subject matters related to humankind, including religion, society, art, sex, and morality. In its applied version, it has largely influenced the clinical approaches to the treatment of mental and behavioral disorders. Two points should be mentioned here: (1) as a form of psychotherapy, psychoanalysis has a few core concepts—the unconscious determination of conscious behavior, the presence of dynamic conflicts between instinctual forces and other social and moral repressive ones, etc.—that have adopted an endless variety of forms maintained by hundreds of schools and groups, but that have raised strong doubts about the therapeutic efficacy and the scientific status of their basic assumptions. In many cases, hermeneutic and philosophical nature of psychoanalysis has been acknowledged, while leaving in suspense its scientific claims. And (2), most of its practitioners have received medical training and licenses, and their techniques are viewed as an essential part of contemporary psychiatry.

3. Gestalt psychology, initiated in Germany in 1912 under the inspiration of German psychologist Max Wertheimer (1886–1943), claimed that all behavior always follows a mental structure in which all stimuli coming from the outside world and from the subject's body form a mental whole that determines open responses to situations. Although it is mainly a theoretical construction, a large number of applications were devised by one of its leading figures, K. Lewin, to solve problems in school and organizational settings. Gestalt theorists put great emphasis on mental variables such as perception, expectations, values, need for achievement, leadership, and group pressure that proved useful in predicting and explaining behavior in different types of social situations. Important developments on group processes were carried out by former Lewin students Dorwin Cartwright, Leo Festinger, and Stanley Schachter, among others; they received much attention in social and clinical settings.

4. Humanistic psychology (sometimes called "the third force") is a reaction to the reductionism imposed by behaviorism and to the irrationality and unconsciousness emphasized by psychoanalysis. It stresses meaning and values as essential determinants of human behavior. Self-fulfillment, life goals, and personal growth are some of the concepts that are at the core of this movement. Born in the United States, it has been largely influenced by European philosophies such as M. Heidegger and J. P. Sartre's existentialism and E. Husserl's phenomenology. Abraham Maslow (1908–1970), Carl Rogers (1902–1987), and Charlotte Bühler (1883–1974) are some of the names associated with it; all of them significantly contributed to psychotherapy, and their insights were taken as bases for a technique of personal development that proved fruitful in the world of management and human relations. Close to these views are some other clinical approaches that also emphasize the role of subjective feelings, cognition, and emotions; these processes deeply influence the subject's mind in its interactions with the world. A. T. Beck's cognitive therapy, A. Ellis's rational emotive therapy, and F. Perls's Gestalt therapy, among others, have widened the spectrum of ways to intervene in psychopathological problems.

5. While these schools largely dominated the U.S. scene, different lines of thought were prevalent in Europe, such as functionalism, a movement with American roots (e.g., William James, John Dewey) but with robust development in Great Britain, Switzerland, and other Western countries. Functionalism stressed the relevance of the "what for" considerations when applied to psychological acts and aptitudes, and mainly focused on life's needs and demands. Industrial psychologist Charles S. Myers (1873–1947) exerted great influence on some of his students; among them was Sir Frederic Bartlett (1886–1969), who was interested in real life problems, and who did significant work on memory and thinking inside a functionalist framework. Swiss functionalist Edouard Claparède applied these views to child and educational psychology, and one of

his collaborators, the Swiss theorist Jean Piaget (1896–1980), deepened those insights and built a major system in epistemology and knowledge, offering a solid view on the biological bases of cognitive development—from reflex action to logical operations—and its sequential stages. As is well known, Piagetian ideas have greatly influenced education in many countries.

6. Last but not least, the Marxian philosophy (historical materialism) that inspired the Soviet regime in Russia (the Soviet Union between 1917 and 1990) also was the basis for a worldview embracing humankind and society that influenced all dimensions of knowledge. Psychology was no exception.

Political theorists such as Karl Marx (1818–1883) and Vladimir Lenin (1870–1924) were considered conceptual leaders in every scientific and cultural field. Human life, guided by the conscious mind, was organized through a process of active interactions with the world, but human actions received their ultimate meaning from the socioeconomic structure in which people lived. Consciousness ultimately depended on the production system dominating society. As a consequence, it stressed the unity between consciousness and activity, and, furthermore, between theory and practice.

Lenin, while denouncing "bourgeois exploitation" in Taylorism and other Western labor theories, gave priority to those studies related to work. In 1927, the All-Russia Psychotechnical Society was established, and in 1931 the 7th International Psychotechnical Conference took place in Moscow. Well-known specialists in educational matters, such as Nadezhda Krupskaya (1869–1939), Isaak N. Spielrein (1891–1936?), and Lev Vygotsky (1896–1934), tried to synthesize the Marxist view with recent Western developments; all of them suffered from prosecution in the days of Jozef Stalin's personality cult in the late 1930s, and for nearly two decades their ideas and works were banned and silenced. During this time, psychology focused on practical (social) activity as causing mental organization through brain structures. Marxian views seemed to fit well with Pavlovian psychophysiology, and theoretical and applied questions became more closely connected. The significant growth of Soviet psychology since the 1950s was based on the cooperation of these two lines of thought. Alexei Leontiev's (1903–1979) idea of activity, largely based on Vygotskyan analyses, spread out to other fields, such as Alexander Luria's (1902–1977) defectology and neuropsychology. The world race for space supremacy, that in the 1960s involved the United States and the USSR, stimulated man–machine research, with outstanding contributions such as Boris Lomov's (1927–1989) industrial psychology and Vladimir Zinchenko's ergonomics, among others.

The end of the Soviet regime in 1990 brought with it the fall of its previous political and philosophical framework, and a new freedom for thought emerged, in which many Western schools and tendencies then flourished.

7. APPLIED PSYCHOLOGY IN THE SECOND HALF OF THE 20TH CENTURY

In the decades that followed World War II, enormous growth took place in psychology that covered most of the facets of human activity and work, in many cases producing interdisciplinary research. Theory also experienced significant changes that have been frequently interpreted as the rise of a new, cognitive paradigm. According to this paradigm, humans should be viewed as information-processing systems, and consciousness, cognitions, affections, and emotions reflect the richness and complexity of such a system. Substantial inspiration for the new paradigm came from applied areas. Swiftly developing computer technology opened new ways to the study of behavior based on the computer model (metaphor) of the brain.

Structure and qualities of behavior referred both to organism and to environment, whose complex interactions had to be analyzed into terms of plans, projects, expectations, and values. Plans and structure of the human operator parallel the software–hardware dimensions of a computer. A vast program launched to examine analogies and possibilities of the new metaphor for psychological research brought up new lines of thought: behavior simulation by computers, artificial intelligence, computer-assisted learning, expert programs of various specialized tasks (e.g., medical diagnosis), and mental activity viewed as the operation of an information-processing system (IPS). Such a system should be viewed as an "open system," in which the operator unit and its environment exchange information signals, energy, materials, and causative interactions. The entire complex (P-E) integrates both person (P) and environment (E) in continuous interaction.

A cursory review of some of the more significant developments that took place in this period are presented here, without any pretension of exhaustiveness.

7.1. Ecological Psychology

The P-E model, as described above, is at the core of an ecological approach to psychology. The environment is mainly viewed in its material aspect, and human behavior is then examined in the light of the physical determinants of its medium and their influences and constraints upon the activity of the subject. Roger Barker (1968), influenced by Lewin's ideas, conceived of the field as a study of objective contexts influencing behavior, a hypothesis full of consequences for other disciplines such as urban and home designs, geriatrics, and child care, among others. At present, environmental psychology, defined as a study of the relationships between people and their physical environments, examines a large range of questions, from coping with disaster or crime and environment to developing positive attitudes toward nature and protecting nature against human destructiveness.

7.2. Traffic Safety Psychology

Closely connected with the preceding, psychology has developed conceptual and technical approaches to behavior of drivers and pedestrians in traffic. Many psychological functions (perceptual, motor, emotions, risk-seeking, and so on) have a strong influence in determining traffic accidents. Although this field is as old as applied research, recent developments in motivation, risky decisions, and accident-prone personality have added new dimensions to it. In several countries, psychologists are now working on the screening of potentially dangerous drivers and in the design of public campaigns for the promotion of safe attitudes in driving situations.

7.3. Organizational Psychology

Organizations are the most frequent social environment for women and men in developed societies. Industries, educative and health institutions, and political organizations are social structures imposing facilities and constraints on their members, which in turn activate their psychological abilities to adapt and to achieve their goals inside those structures. As a matter of fact, this denomination has nearly come to represent what industrial psychology was covering some decades ago (e.g., personnel selection, human relations), stressing its structural approach to these types of problems.

Some significant lines of thought on this field that developed in the early decades of the century were noted earlier. In the second half of the century, organizational psychology has grown, and in some cases it has dominated the sociotechnological approach. Other groups stressed the relevance of humanistic aspects and of combining the needs of both the individual and the organization, trying to arrive at a situation of balanced forces contributing to the whole structure and viewing organizational behavior as based on bounded rationality.

Organizations operate in a changing world, and need to maintain their adjustment to the environment through changes taking place inside and outside their own boundaries. A vast technology has been developed to provide them with techniques and means for their development. Organizations, through changes in structure, culture, competence, and processes, aim at improving their problem solving and conflict resolution and their own clime and efficiency. Far from the old lines of rationalist scientific management, the organizational development, rooted in the action research theory designed by K. Lewin in the 1940s, operates with psychological variables through micro and macro techniques, including program assessment and redesigning of structures; its topics have become central in the research field.

Special cases of organizations are armies, churches, national administrations, and so on. These large social structures generate enough questions to become the particular object of psychological specialties covering them. The use of telematics and new information and communication technologies applied to organizations have also raised new demands on workers' formation and capacities, an interdisciplinary field that many psychologists are now entering.

Other fields, such as the large one of sport behavior or military psychology, can also be mentioned in this regard. In the former, psychologists have proved to be efficient technicians in dealing with variables that determine high quality sports performances (e.g., achievement motivation, emotional and attention control). The latter includes the study of specific organizations (armies) that were among the first clients of psychology during wartime. In this field, there is clear overlap with similar areas such as engineering psychology and equipment design and ergonomics. In recent times, emphasis has turned to relevant personality dimensions like authoritarianism, leadership, conformity, and prejudice that set the basis for social conflicts. During the so-called Cold War between Western democracies and the Soviet world, some topics also gained relevance, such as the "war of nerves," national morals, and conflict simulation.

In their approach to such complex problems, some theorists have explored the multiplicity of determinants that influence human action, trying to understand the cues and factors that operate in effective choices and decisions; such insights have enlarged the well-known Lewinian "action theory," which tries to strengthen theory and practice bonds.

7.4. Health Psychology

This branch, which is closely related to the clinical field, has rapidly developed since the 1980s. It concentrates on all psychological questions related to health maintenance, illness prevention, and patient–practioner interactions in treating diseases. This specialty deals not only with problems of the individual, but also with others of an organizational nature linked to health care systems and health policy. Prevention, adherence to doctor's instructions, and people's healthy life habits are now among the goals of psychologists working in this field.

Connected to this field is contemporary psychogerontology, a specialty that emerged from studies on lifespan developmental psychology and turned into one of the most active areas of research and social priority. In many Western countries, the elderly occupy one of the largest segment of the population, and retirement arrives at an age still able for productive activity. Leisure, social deprivation, loss of power and influence, depressive moods, and psychological losses became relevant targets for psychological intervention.

A special chapter has recently developed and is receiving great attention from society: stress research. Growing evidence shows a causal influence of psychological factors upon the immune system of organisms. Stress research largely deals with the relationship between psychological and organic variables and stressing experiences. The style of coping with negative effects seems not unrelated to serious diseases such as cancer, heart attacks, and AIDS, and with those organic processes connected with the end of such processes. For instance, much literature has been produced on A personality types, considered by some researchers to be risk factors of many serious diseases, mediating some changes in the bodily immune system.

It is also the important and complex role assumed by psychologists working on drug-dependence assistance, one of the most serious problems in contemporary society. Abuse and addiction therapies demand both physical and psychological attention, and patients require social support and advice for changing their attitudes toward dangerous substances and to adhere to treatment.

Finally, the clinical field is also neuropsychology. This is one of the areas that has experienced extraordinary developments in recent times. It is usually described as the study of the brain machinery underlying psychological functions in patients suffering from brain damage, carried out for practical purposes of rehabilitation. Its conceptual roots can be found in the works of Alexander R. Luria (1902–1977) and Donald O. Hebb (1904–1985). Technical advances in brain diagnosis through direct imagery (positron-emission tomography and the like) have combined with systematic testing of behavioral effects of damage through batteries such as the Halstead-Reitan Neuropsychological test battery, the Luria/Christensen test battery, and many others. Cognitive and emotional deficits and perturbations, while significant for the clinician, are also relevant for the general theorist, who tries to combine the experimental cognitive approach with new neuropsychological information in order to deepen the knowledge on the mind–brain relationship. This is clearly an interdisciplinary field that needs the joint labor of neurologists, psychologists, and other scientists.

7.5. Law Psychology

Since its very beginning, applied psychology has been charged with the task of establishing the degree of credibility of eyewitness testimonies presented in court. Karl Marbe (1869–1953) studied child charges of sexual abuse, and, interested in the topic, wrote a pioneering handbook on forensic psychology. U.S. clinics for offenders' rehabilitation created by Healey have been mentioned previously. Since then, questions of deception in testimonies, jury selection, legal evidence in court, deterrence from crime, rehabilitation of offenders, and assessment of convict personalities, among others, are under scrutiny and discussion. In the United States, an important court sentence that accepted the psychologist as an expert on mental states of defendants in court strongly promoted the specialty.

7.6. Other Specialties

Human behavior offers endless dimensions and aspects that may be considered by the psychologist. Apart from those mentioned above, other specialties that are now flourishing in developed countries include community

psychology, economic psychology, and political psychology. Psychological care is increasingly a right for people living in developed countries. These provide their citizens with a welfare system that offers security and support when individuals and groups have to cope with stressful events, which frequently include psychological connotations. Problems of family life, such as abuse and neglect of children and women, victimology, and crime and drug prevention, require planning that should include psychological assessment and measures in order to be effective.

Modern societies rely on a large economic system, that through tax policies and contributions supports the major needs of administrative and social services that are in charge of the state. While societies for centuries employed constrictive measures, more attractive ones have been introduced by democratic regimes, which try to persuade people to act lawfully by spontaneous motivation and reflection instead of by compulsion. Individuals who are forced to participate in supporting the state budget, and who are active members of the market system, are basically guided by their interests, beliefs, and fear of penal enforcement. Techniques employed to increase efficiency and motivation in business in the early days of applied psychology by W. D. Scott and other pioneers have recently been transferred and adapted to social policy by governments and social agencies. Economic psychology today covers a wide variety of topics, including behavior in market settings, taxes, investing, risk assumption, bargaining, publicity, and consumer behaviors.

8. GENERAL TRENDS

Most psychological professionals work in the field of applied psychology, and the field produces a large portion of the research in psychology. While it is usual to think that it deals with the mere application of some existing principles to certain circumstances, practical intervention frequently gives birth to new concepts and techniques in a continuous interaction among proposed goals, effective applications, and evaluations.

Applied psychology systematically uses evaluation techniques in order to determine the initial problem, to measure the changes produced in it by psychological intervention, and to assess the resulting costs and benefits. According to Gilgen and Gilgen (1987), "Whatever the country . . . it seems clear that support for academic research psychology is largely dependent on how successful psychologists are at demonstrating their usefulness to society."

Modern technological devices (e.g., complex computer-assisted simulation, modeling and data analyses) have been introduced in all social organizations and working groups, and psychologists have assumed the demands created by that technology, and have incorporated them in their own practice. This has resulted in the task of training and recycling specialists as well as a redesign of groups in view of the new goals to be reached.

In every field, interventions have generated the creation of micro-level models that try to fit with concrete data, and whose roots on a general view of mind and behavior are loose and flexible. Assessment runs parallel with intervention, in order to introduce the needed corrections to the planned operation.

Interdisciplinarity has become dominant, and psychologists are approaching the views and methods of those groups with which they have to cooperate. At the present, tremendous development in fields such as genetics, bioengineering, and computer sciences demand further and deeper interaction with psychology, for both theoretical and practical reasons. The new situation demands specialization and competence in many scientific disciplines, but at the same time, all progress will depend upon the ability to connect these new ideas to the questions raised by the understanding of the human mind.

In recent decades, psychologists involved in professional practice clearly outnumber those oriented toward academic research. Their relative weight has turned upside-down the situation that existed in the early days. In addition to the debates on scientific topics, new questions are demanding solutions from professionals: the implementation of adequate training (in concepts and skills) and a qualifications for licensing in different specialties; the adaptation of some minimal technical standards to the peculiarities and idiosyncrasies of the different societies and cultures in which psychology is now applied; communication between academics and the practical world, benefiting from recent findings; the creation of mechanisms for accumulating and disseminating the growing body of knowledge in useful and standardized ways; the updating of ethical codes in view of the multiplicity of communication systems and computational technologies; and the increasing penetration of psychology in different aspects and dimensions of modern life.

Training for a professional career in psychology increasingly requires a specialization in concrete techniques that run the gamut of related topics. In many cases, this is deferred until the postgraduate level.

International cooperation provides additional possibilities for effective work. Associations such as the International Association of Applied Psychology, the European Federation of Professional Psychologists (recently renamed the European Federation of Psychological Associations), and the American Psychological Association are considering the requirements to become a well-trained professional in the various branches of psychology, and the technical and conceptual requirements required by the growing complex interactions between this professional and its clients.

See Also the Following Articles

Clinical Assessment ■ Cross-Cultural Psychology, Overview ■ Cultural Psychology ■ Environmental Psychology, Overview ■ Health Psychology, Cross-Cultural ■ Industrial–Organizational Psychology across Cultures

Further Reading

Adler, N., & Matthews, K. (1994). Health psychology: Why do some people get sick and some stay well? *Annual Review of Psychology*, *45*, 229–259.

Ardila, R. (1978). *La profesión del psicólogo*. México, Trillas.

Baker, D. B., & Benjamin, L. T., Jr. (2000). The affirmation of the scientist-practitioner. A look back at Boulder. *American Psychologist*, *55*, 241–246.

Baltes, P. B., & Baltes, M. M. (1990). *Successful aging*. New York: Cambridge Univ. Press.

Barker, R. (1968). *Ecological psychology*. Stanford: Stanford Univ. Press.

Baumgarten, F. (1957). *Examenes de aptitud profesional. Teoría y práctica*. Barcelona, Labor.

Birren, J. E., & Schaie, K. W. (Eds.) (1900). *Handbook of aging and the individual*. Chicago, Chicago Univ. Press.

Boden, M. (1988). *Computer models of mind*. New York, Cambridge Univ. Press.

Brigham, C. (1923). *A study of American intelligence*. Princeton, NJ: Princeton Univ. Press.

Brozek, J. (1973). Soviet psychology, In Marx, M. H., & Hillix, W. A. (Eds.), *Systems and theories in psychology* (2nd ed., pp. 521–548). New York: McGraw-Hill.

Buckley, K. W. (1989). *Mechanical man: John Broadus Watson and the beginnings of behaviorism*. New York: Guilford Press.

Bühler, C., & Allen, M. (1972). *Introduction to humanistic psychology*. Monterey, CA: Brooks & Cole.

Carpintero, H. (2001). The psychology of aging in historical perspective, In Leon-Carrión, J., & Giannini, M. (Eds.). *Behavioral neurology in the elderly* (pp. 3–22). Boca Raton, FL: CRC Press.

Chanderssais, C. (1961). *La psicología en las fuerzas armadas*. Buenos Aires: Kapelusz.

Cohen, S., & Herbert, T. (1996). Health psychology: Psychological factors and physical disease from the perspective of human psychoneuroimmunology. *Annual Review of Psychology*, *47*, 113–142.

Compayré, G. (1898). *Psicología teórica y práctica aplicada a la educación*. Paris: Lib. V. C. Bouret.

Dorsch, F. (1963). *Geschichte und Probleme der angewandten Psychologie*. Bern: Hans Huber.

Du Bois, P. (1970). *A history of psychological testing*. Boston: Allyn and Bacon.

Eysenck, H. J. (1971). *Race, intelligence and education*. London: Temple Smith.

Fernandez-Ballesteros, R. (1997). *Psicologia del envejecimiento: crecimiento y declive*. Madrid: Univ. Autonoma de Madrid.

Fine, R. (1990). *The history of psychoanalysis* (2nd ed.). New York: Continuum.

French, W. L., & Bell, C. H. (1978). *Organizational development* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Geuter, U. (1992). *The professionalization of psychology in Nazi Germany*. Cambridge, UK: Cambridge Univ. Press.

Gilgen, A. R. (1982). *American psychology since World War II: A profile of the discipline*. Westport, CT: Greenwood Press.

Gilgen, A. R., & Gilgen, C. K. (1987). *International handbook of psychology*. New York: Greenwood Press.

Gould, S. J. (1981). *The mismeasure of man*. New York: Norton.

Gundlach, H. (1998). An outline of the history of the IAAP, In *Applied Psychology: The First-Thirteenth Congress Proceedings of the International Association of Applied Psychology*, Vol. I (pp. 1–24). London: Routledge.

Hearnshaw, L. (1964). *A short history of British psychology*. London: Methuen.

Herrnstein, R., & Murray, C. (1994). *The Bell Curve: Intelligence and class structure in American life*. New York: The Free Press.

Holtzman, W. H. (1992). Health psychology, In Rosenzweig, M. (Ed.,) *International Psychological Science* (pp. 199–226). Washington, DC: APA-IUPsyS.

IAREP (1997). *IAREP XXII International Colloquium of Economic Psychology*. Valencia, Promolibro.

Jaeger, S., & Staeuble, I., (1981). *die psychotechnik und ihre gesellschaftlichen entwicklungsbedingungen, in Die Psychologie des 20* (pp. 53–95). Jahrhunderts: Zurich.

James, W. (1899/1962). *Talks to teachers on psychology and to students on some of life's ideals*. New York: Dover.

Jensen, A. J. (1973). *Genetics and education*. New York: Harper & Row.

Johnson-Laird, P. N. (1988). *The computer and the mind. An introduction to cognitive science*. Glasgow: Collins & Sons.

la Vaissière, J. (1915). *Psychologie pédagogique*. Paris: Beauchesne.

Leon-Carrion, J. (Ed.) (1997). *Neuropsychological rehabilitation: Fundamentals, innovations, and directions*. Delray Beach, FL: GR/St. Lucie Press.

Leonova, A. (1995). Applied psychology in the European North-East: Historical traditions, stages of development and current perspectives. In Jaeger, S., *et al*. (Eds.), *Psychologie Im Sozialkulturellen Wandeln.-Kontinuitäten und Diskontinuitäten* (pp. 186–193). Frankfurt: Peter Lang.

Lévy-Leboyer, C. (1980). *Psychologie et environnement*. Paris: PUF.

Lewin, K. (1948). *Resolving social conflicts*. New York: Harper & Row.

Lopes, L. (1994). Psychology and economics. *Annual Review of Psychology*, *45*, 197–227.

Lunt, I. (1999). The professionalization of psychology in Europe. *European Psychology*, *4*, 240–247.

Marbe, K. (1913). *Grundzüge der forensischen Psychologie*. München: Beck.

Matarazzo, J. (1980). Behavioral health and behavioral medicine: Frontiers for a new health psychology. *American Psychologist*, *35*, 807–817.

Monahan, J., & Loftus, E. (1982). The psychology of law. *Annual Review of Psychology*, *33*, 441–475.

Montoro, L., *et al*. (1995). *Seguridad vial: Del factor humano a las nuevas tecnologías*. Madrid: Síntesis.

Näätänen, R., & Summala, H. (1976). *Road-user behavior and traffic accidents*. Amsterdam: North Holland.

Otto, R. K., & Heilbrun, K. (2002). The practice of forensic psychology: A look toward the future in light of the past. *American Psychologist*, *57*, 5–18.

Peiró, J. M. (1983). *Psicología de la organización*. Madrid: UNED.

Perusia, F. (1994). *Psicologo: Storia e attualitá di una professione scientifica*. Torino: Bollati Boringhieri.

Petrovsky, A. (1990). *Psychology in the Soviet Union*. Moscow, Progress.

Raimy, V. C. (Ed.) (1950). *Training in clinical psychology*. New York: Prentice Hall.

Reisman, J. M. (1991). *A history of clinical psychology* (2nd ed.). New York: Hemisphere.

Rodriguez Marin, J. (1995). *Psicología social de la salud*. Madrid: Síntesis.

Rüegsegger, R. (1986). *Die Geschichte der angewandten Psychologie, 1900–1940*. Bern: Huber.

Sexton, V. S., & Hogan, J. D. (1992). *International psychology. Views from around the world*. Lincoln, NE: Univ. of Nebraska Press.

Sollier, P., & Drabs, J. (1935). *La psychotechnique*. Bruxelles: Comité Central Industriel de Belgique.

Sprung, L., & Schönpflug, W. (Eds.). (2003). *Zur Geschichte der Psychologie in Berlin* (2nd ed.). Frankfurt: Peter Lang.

Stokols, D. (1978). Environmental psychology. *Annual Review of Psychology*, *29*, 253–295.

Sully, J. (1900). *Psicología pedagógica*. New York: Appleton (originally *The teacher's handbook of psychology*. London: Longmans Green, 1886).

Sundstrom, E., *et al*. (1996). Environmental psychology, 1989–1994. *Annual Review of Psychology*, *47*, 485–512.

Thorndike, E. L. (1910). *Educational psychology* (2nd ed.). New York: Columbia Univ.

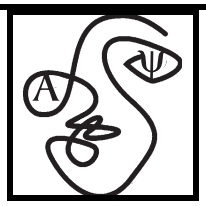
Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in one mental function upon efficiency of other functions. *Psychology Review*, *8*, 247–261, 384–395, 553–564.

Wilpert, B. (1995). Organizational behavior. *Annual Review of Psychology*, *46*, 59–90.

Wilpert, B., & Lunt, I. (1998). *World study of applied psychology*. Unpublished report to IAAP.

Wrightsman, L. (2001). *Forensic psychology*. Belmont, WA: Wadsworth.

Zilboorg, G., & Henry, GW., (1941). *A history of medical psychology*. New York: Norton.



Holland's Theory (Vocational Personality Types)

W. Bruce Walsh

The Ohio State University, Columbus, Ohio, USA

1. The Personality Types
2. The Model Environments
3. Person–Environment Congruence
4. Consistency, Differentiation, and Identity
5. Research
6. Summary
   Further Reading

GLOSSARY

congruence The degree of fit between personality type and work environment.
consistency The degree of relatedness between personality types or between environmental models.
differentiation The degree to which a person or environment is well defined.
identity The clarity and stability of a person's interests, abilities, and goals.
type A cluster of personal attributes.

John Holland's theory of personality types and model environments is a theory of vocational psychology, personality, and person–environment psychology. Holland's theory, as viewed by Helson and Mitchell, is tough, practical, and useful. The theory is operational, and there is much data to support it. In essence, the theory suggests that behavior is a function of the complementary match between an individual's personality style and the psychological environment. Holland suggests that individuals enter environments because of their personalities and remain in those environments because of the reinforcements and satisfactions obtained through interpersonal interactions in the environments.

Holland's theory assumes that behavior is a function of personality and social environment. The theory emphasizes interests and personality since to Holland the choice of an occupation is an expression of personality; thus, interest inventories are personality inventories. The environmental component of the theory is linked to the notion that people in a particular environment tend to have similar personalities and histories of personal development. Because people in a given environment have similar personalities, Holland believes they tend to cope in similar ways. In general, Holland holds that complementary person–environment interactions are reinforcing and satisfying and contribute to job stability, work quality, and psychological well-being.

1. THE PERSONALITY TYPES

Holland attempts to explain behavior by using a few well-defined ideas. These assumptions of his theory

©2004 Elsevier Inc.
All rights reserved.

tend to be fairly practical, straightforward, and operational in terms of data-collection procedures. The first of these assumptions is that people learn to be one or more personality types. A type is defined as a cluster of personal attributes that may be used to assess the person. Holland's six personality types are described as follows:

1. The realistic (R) type likes realistic jobs such as automobile mechanic, farmer, or electrician. R's have mechanical abilities and tend to be conforming, honest, materialistic, natural, persistent, practical, modest, and stable.
2. The investigative (I) type likes investigative jobs such as chemist, physicist, biologist, or geologist. I's have mathematical and scientific ability and are described as analytical, independent, intellectual, introverted, methodological, and rational.
3. The artistic (A) type likes artistic jobs such as musician, writer, or actor. A's have writing, music, and/or artistic abilities and are described as complicated, emotional, expressive, imaginative, impulsive, nonconforming, and original.
4. The social (S) type likes social jobs such as teacher, counseling psychologist, clinical psychologist, or speech therapist. S's have social skills and are described as cooperative, friendly, helpful, insightful, responsible, sociable, and understanding.
5. The enterprising (E) type likes enterprising jobs such as salesperson, manager, or business executive. E's have leadership and verbal abilities and are described as ambitious, domineering, pleasure seeking, self-confident, and sociable.
6. The conventional (C) type likes conventional jobs such as bookkeeper, banker, or tax expert. C's have clerical and arithmetic ability and are described as conforming, conscientious, orderly, persistent, practical, and self-controlled.

A person's resemblance to each of the personality types is assessed by the use of the Vocational Preference Inventory, the Self-Directed Search, and the Strong Interest Inventory. All of these inventories are objective in nature and collect interval kinds of data that may be used in a variety of research settings. Much of the research on Holland's theory has used one of these instruments in a field research setting.

2. THE MODEL ENVIRONMENTS

A second idea used by Holland to explain behavior is that environments in which people live and work may be characterized by their resemblance to one or more model environments. Holland suggests six model environments corresponding to the six personality types and believes that people tend to move toward environments congruent with their personality types: Realistic types tend to move toward realistic environments, enterprising types tend to move toward enterprising environments, and so forth. To assess environments, Holland and Astin developed the Environmental Assessment Technique. This assessment technique simply makes a count or a census of the number of different personality types existing or interacting in a given environment. The personality type occurring most frequently in a given environment is thought to be dominant in the environment and to have a psychological influence in that specific environment. In summary, Holland defined the environment psychologically according to the people who were actually interacting in it, or, as he would say, people define the environment. Thus, although the environment is determined by a count, the influence is thought to be psychological in nature.

3. PERSON–ENVIRONMENT CONGRUENCE

A third idea used by Holland is that person–environment congruence tends to be associated with satisfaction, productivity, personal stability, vocational stability, and well-being. The thought is that, for example, a realistic type in a realistic environment would tend to be more productive, more stable, more vocationally satisfied, and psychologically healthier. Complementary person–environment links are reinforcing and satisfying. Uncomplementary or incongruent person–environment interactions are punishing and contribute to change. This is not to say that congruence is good and incongruence is bad but suggests that a state of incongruence may be part of the learning and development process. Thus, there seems to be periods of incongruence in an individual's life span that need to be worked through in order to pursue personal and vocational goals and well-being.

Based on his six personality–environment types, Holland further suggests that the relations among types, among environments, and between types and environments tend to be hexagonal in nature (Fig. 1). Types and environments that are in closest proximity in Fig. 1 are more psychologically related. Types and environments that are further removed from one another are more psychologically different. For example, an artistic person

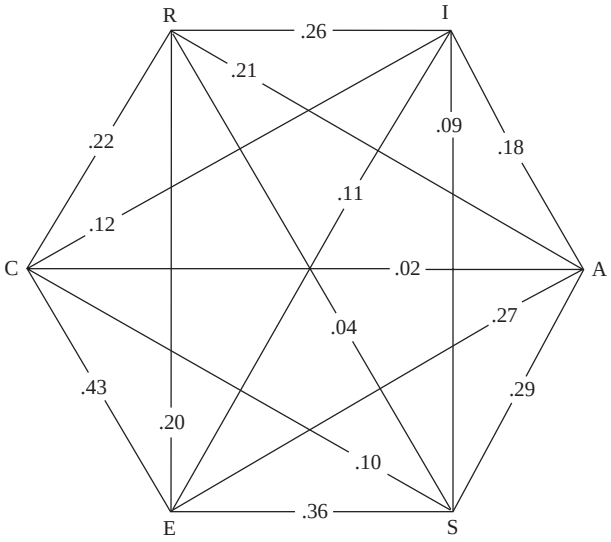


FIGURE 1 Hexagonal model for interpreting inter- and intraclass relationships. Correlations are between summary scale scores for females (n = 1600) in the 1994 normative sample. Reproduced by special permission of the publisher, from the *Self-Directed Search Technical Manual* by John L. Holland, Ph.D. Copyright 1985, 1987, 1994, by Psychological Assessment Resources, Inc. Further reproduction is prohibited without permission of PAR, Inc.

in an artistic environment is in a more congruent situation than an artistic person in a social environment. An artistic person in a conventional environment is in the most incongruent situation possible. In this way, the hexagonal model may be used to obtain and estimate different degrees or levels of person–environment congruence. The hexagonal model for the six personality types and environments has also been used to identify individuals in vocational or personal conflicts, compare different tests and inventories, organize occupational environments, and guide the layout of the Strong Interest Inventory.

4. CONSISTENCY, DIFFERENTIATION, AND IDENTITY

The previous assumptions, as noted by Holland, are extended by several secondary assumptions that may be applied to both persons and environments. The secondary assumptions include consistency, differentiation, and identity. Consistency is the degree of relatedness between personality types or between environmental models, as noted previously. Some pairs of types and environments are more closely related than others. Differentiation refers to the degree to which a person or an environment is well defined. For example, a person closely resembling a single type is clearly differentiated (well defined). Identity is primarily concerned with estimating the clarity and stability of a person's vocational identity. High identity suggests a limited set of consistent and explicit goals.

5. RESEARCH

In general, research testing the theory (more than 500 studies) reported by Holland and Walsh and Holland clearly indicates that individuals tend to choose and enter environments consistent with their personality types. For example, enterprising types tend to enter and remain in enterprising kinds of environments. Furthermore, evidence suggests that congruent person–environment interactions are conducive to personal and vocational stability and satisfaction and, to a lesser extent, actual achievement. Specifically, the pattern of findings suggests to some extent that person–environment congruence is related to measures of job satisfaction and stability, job involvement, work quality, productivity, and well-being but not necessarily to measures of decision making, sociability, and problem-solving ability. In addition, meaningful data about adolescents and adults imply that person–environment congruence and job satisfaction may increase with age. Thus, it seems that people in congruent work situations will probably change little since their expression of interests, abilities, and personality is rewarded. Persons in incongruent work environments will probably change the most since they tend to be ignored and punished. This work, then, indicates that students and workers who tend to be functioning in environments congruent with their personal characteristics are probably psychologically healthier and more satisfied than people in incongruent situations.

6. SUMMARY

In Holland's theory, behavior is viewed as a function of person–environment interactions, and from this perspective person–environment congruence becomes a very important concept. Thus, Holland suggests, and the data tend to support, the notion that people in environments congruent with their personality types tend to be psychologically healthier, more satisfied,

and more productive than people in incongruent environments. In other words, person–environment congruence tends to facilitate personal growth, self-acceptance, and life purpose.

See Also the Following Articles

Person–Environment Fit ■ Traits ■ Vocational Interests ■ Work Adjustment

Further Reading

Astin, A. W., & Holland, J. L. (1961). The environmental assessment technique: A way to measure college environments. *Journal of Educational Psychology*, *52*, 308–316.

Harmon, L. W., Hansen, J. C., Borgen, F. H., & Hammer, A. L. (1994). *Strong Interest Inventory: Applications and technical guide*. Stanford, CA: Stanford University Press.

Helson, R., & Mitchell, V. (1978). Personality. In M. R. Rosenzweig, & L. W. Porter (Eds.), *Annual review of psychology* (pp. 555–585). Palo Alto, CA: Annual Reviews.

Holland, J. L. (1985). *Vocational Preference Inventory*. Odessa, FL: Psychological Assessment Resources.

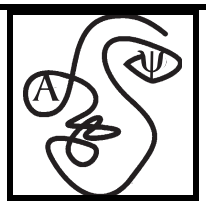
Holland, J. L. (1994). *Self-Directed Search (SDS) Form R*. Odessa, FL: Psychological Assessment Resources.

Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3rd ed.). Odessa, FL: Psychological Assessment Resources.

Spokane, A. R. (1985). A review of research on person–environment congruence in Holland's theory of careers. *Journal of Vocational Behavior*, *26*, 306–343.

Spokane, A. R., Meir, E. I., & Catalano, M. (2000). Person–environment congruence and Holland's theory: A review and reconsideration. *Journal of Vocational Behavior*, *57*, 137–178.

Walsh, W. B., & Holland, J. L. (1992). A theory of personality types and work environments. In W. B. Walsh, K. H. Craik, & R. H. Price (Eds.), *Person–environment psychology*. Mahwah, NJ: Erlbaum.



Home–School Collaboration

Richard J. Cowan
Kent State University, Kent, Ohio, USA

Susan M. Swearer and Susan M. Sheridan
University of Nebraska, Lincoln, Nebraska, USA

1. Home–School Collaboration Defined
2. Background and History of Home–School Collaboration
3. Factors Influencing Home–School Collaboration
4. Approaches to Fostering and Maintaining Home–School Collaboration
5. Conclusion
Further Reading

GLOSSARY

collaboration Working together toward a common goal or set of goals.

educators General education instructors, special education instructors, teaching aides, classroom volunteers, administrators, support staff, school psychologists, and any other persons within a school setting who contribute to the education of students within that environment.

parent The adult who is responsible for the care and well-being of the child (e.g., biological parent, caregiver, legal guardian).

partnership A coequal interdependent relationship that is established and developed over time, with the primary purpose of working together over an extended period of time toward a mutually determined set of goals and objectives.

system A particular environment with an emphasis on specific components of the environment, including people, physical objects, placement of people and objects, physical arrangement, and interactions among these components (e.g., home system; school system).

Home–school collaboration is a reciprocal dynamic process that occurs between at least one parent (or guardian) and at least one individual within the school system (e.g., educators, administrators, psychologists) who share in decision making regarding mutually determined goals and solutions related to a student for whom all parties share interest and responsibility. Home–school collaboration may occur at three levels: system (i.e., between the home and school settings), school/classroom, and individual. With all participants (e.g., parents, educators) offering their unique roles and contributions, the collaboration process is guided by the primary emphasis on mutually determined academic and behavioral goals serving as the standard for progress toward overall outcomes.

1. HOME–SCHOOL COLLABORATION DEFINED

A common definition of the term "collaboration" is to work together toward a common goal or set of goals. As applied to relationships among parents, educators, and the systems in which they operate, the term has a rich meaning. Driven by academic, cultural, and political influences, the term "home–school collaboration" is a reciprocal dynamic process that occurs among systems (e.g., families, communities, partnerships), schools/classrooms, and/or individuals (e.g., parents, educators, administrators, psychologists) who share in decision making toward common goals and solutions related to students. Inherent in this definition is the notion that parents and educators involved in collaboration pool

©2004 Elsevier Inc.
All rights reserved.

their resources to create a cooperative interdependent relationship. For example, parents offer what they know about their children's academic and behavioral strengths and limitations in the home and other non-school settings, whereas teachers offer their knowledge of students' relative strengths and limitations in the classroom and other educational settings. The collaboration process is guided by a primary emphasis on specific cooperatively predetermined outcomes for students, with mutually established academic and behavioral goals.

2. BACKGROUND AND HISTORY OF HOME–SCHOOL COLLABORATION

2.1. Traditional Family Involvement in Education

During colonial times and throughout the early 1900s, parents had very little interaction with the educational system. In 1995, Welch and Sheridan noted that during that time in educational history, children were sent to school to learn academic skills, and parents were seen as responsible for educating their children regarding moral principles in the home setting. Parents entrusted the education of their children to school personnel and expressed little interest in participating in their children's education. These expectations and roles resulted in educators viewing the school as their territory, and parents who tried to become an active part of the school setting were viewed as intrusive. In response to this situation, the National Parent–Teacher Association (PTA) was founded on February 17, 1897, in Washington, D.C., with the goal of having a PTA in every school. Today, there are more than 6.5 million members of the National PTA, making it the largest child advocacy organization in the United States.

Welch and Sheridan reported that parents continued to assume a relatively "hands-off" approach to education until the late 1960s and early 1970s, when economic and political factors induced parents to change their views regarding their involvement in education. Fueled by economic situations that forced society to reconsider how money was being spent on education, as well as by a political climate in which legislation was identifying and defining specific roles for parents in their children's education, parents became more active in education. The National PTA continued to be a strong advocacy voice for parent involvement in education. In 1994, the federal government enacted "Goals 2000," and in 1997, the National PTA released its National Standards for Parent/Family Involvement Programs. However, although parents advocated for participation in education, the educational system was not prepared for this type of parental involvement. Despite the efforts of the National PTA, when parents participated in education, their involvement was limited in both scope and depth. Unfortunately, parental involvement in education has continued to be limited. Many parents feel as though they are outsiders in their children's school. Merely opening the doors for parents to participate at an entry level is not enough. Epstein has written extensively about the need to develop family–school–community partnerships to enhance student performance. These partnerships reflect an important process that includes all families. The next subsection describes a philosophy to guide parental involvement in education: the home–school partnership model for working with families.

2.2. Partnering With Families in Education: A Modern Approach to Collaboration

The definition of home–school collaboration provided earlier is derived in part from the characteristics of a partnership philosophy for interacting with families in education. This philosophy represents the current standard for how parents and families can be included as a member of the educational system. A partnership philosophy regarding home–school collaboration recognizes that parents are invaluable members of the educational team. Just as educators and administrators are viewed as possessing educational information and expertise, parents are viewed as experts possessing equally critical information regarding the home system. When schools adopt this attitude, they are more likely to actively pursue meaningful parental participation in education. If parents and educators adopt this philosophy, it affords parents equal status with the school regarding critical activities and decision making related to their children's education.

A primary difference between home–school collaboration and home–school partnership is that home–school collaboration implies a process related to a specific goal or set of goals that may be relatively short term (but not necessarily), whereas home–school partnership implies a long-term, ever-evolving relationship between parents and members of the school setting extending beyond time-limited problem solving and goal achievement. A partnership allows parents and educators to work together to improve the educational

system such that it is possible for them to work toward preventing future problems, meaning that all students benefit from the partnership. Another characteristic of a partnership is that individuals learn from one another and pool together expertise and skills in support of a common focus, that is, improved education for all students. Goals and objectives remain at the heart of interactions between parents and educators as they form educational partnerships; however, with partnerships, these goals are likely to be more long term and more encompassing (e.g., creating a safe and diverse atmosphere that addresses the needs of all students). In addition, the quality of the partnership will depend heavily on the type of communication that occurs between parents and educators. Maintaining open communication is likely to make a meaningful contribution to a healthy balanced partnership.

When schools invest in building partnerships by seeking meaningful ways in which to involve parents, many can be considerably enriched. For example, instead of a traditional one-directional approach to parent–teacher conferences, a teacher might approach this event by asking for parental input regarding behavior patterns across nonschool settings and ask for ideas regarding potential solutions. Furthermore, if parents and teachers interact regularly outside of the designated parent–teacher conference time (e.g., through home–school notes, at school activities), conference time might be used more constructively to address goals and solutions that require face-to-face communication. A partnership is not formed and developed during one interaction or activity; rather, it is established and cultivated over time through multiple interactions during which parents and educators come to embrace the attitude and belief that meaningful relationships form the foundation of in-depth, long-term collaboration.

2.3. Potential Benefits of Home–School Partnerships

There are many potential benefits of home–school partnership. For example, collaboration through partnership can enhance communication between parents and educators. In addition, home–school relationships provide an opportunity for parents and educators to better understand what occurs across settings, and this in turn allows for a greater overall understanding of children's backgrounds and current levels of functioning. Establishing a home–school partnership also allows for "joint ownership" of problems and concerns as well as joint commitment toward educational and behavioral goals for students. Parents and teachers may come to feel supported by each other as they work toward mutually determined goals. Working together also tends to improve relationships between teachers and parents as they begin to understand each other's perceptions regarding children's behavior across settings. In addition to these potential benefits for parents and educators, students may benefit from the process. For example, they may develop better attitudes toward school, improve their performance both in and out of the classroom (e.g., demonstrate higher rates of work completion at school while at the same time demonstrating an increase in homework completion), get along better with teachers and other school-based individuals, and improve specific academic skills (e.g., reading fluency). As with any relationship, the benefits of a partnership might not be immediate, and it will likely take time for all participants to yield such benefits.

3. FACTORS INFLUENCING HOME–SCHOOL COLLABORATION

The educational system in the United States has been heavily influenced by a systems-based approach to understanding and intervening with regard to the educational process. This approach is grounded in ecological theory. Based on the work of Bronfrenbrenner, ecological theory is concerned with the interaction between an individual and various contextual systems: the microsytem, mesosystem, exosystem, and macrosystem. The microsystem is concerned with the individual and his or her immediate environment (e.g., a child interacting in a particular classroom), the mesosystem involves the interrelation between major systems in the individual's life (e.g., between home and school), the exosystem is concerned with environments not directly related to the individual but still influencing his or her life (e.g., a parent's workplace), and the macrosystem includes overall cultural or subcultural patterns and influences (e.g., policies, federal and state legislation, national and global economic factors). This framework illustrates how systems influence, and are influenced by, one another, stressing that individuals do not exist in isolation; rather, each individual's interactions across and within systems are multifaceted and multidetermined over time. Home–school collaboration is influenced by multiple factors, including (but not limited to) educational legislation, family characteristics,

school-based variables, and community-related influences. Factors across these systems are not mutually exclusive and are discussed in the following subsections.

3.1. Educational Legislation

Education in the United States has forever been changed due to federal legislation that ensures the right to a free and appropriate public education for all individuals, mandates the individualization of services for students with special needs, incorporates the provision of educational services in the least restrictive environment, provides guidelines for educational planning and programming, specifies procedures for identifying students with special needs, highlights roles and responsibilities for local educational agencies, and mandates specific roles for parents in education. Such legislation has led parents, educators, politicians, and others concerned with education to consider how best to foster relationships between parents and educators. Furthermore, it has changed the ways in which parents are involved in their children's education. The following subsections discuss major legislative and political factors affecting how parents and educators operationalize and designate roles and responsibilities in education.

3.1.1. Individuals with Disabilities Education Act

The Individuals with Disabilities Education Act (IDEA) is the latest manifestation of the public education law called the Education for All Handicapped Children Act (EAHCA, also referred to as Public Law 94-142). A major emphasis of the 1997 reauthorization of IDEA is the inclusion of specific requirements regarding the active recruitment and participation of parents throughout the evaluation process. Specifically, school personnel are now required to (a) notify parents in advance to ensure that they will have the opportunity to attend each meeting related to their children's educational programming, (b) schedule meetings at a mutually determined time and place (even in the home if necessary and appropriate), and (c) arrange other means of including parents if they are unable to attend the meeting (e.g., teleconferencing).

In addition to ensuring that parents are able to actively participate in meetings related to their children's educational programming, IDEA mandates and outlines several parental rights and responsibilities. Specifically, parents are entitled to input during evaluation, eligibility, and placement. That is, in addition to contributing to the pool of information that will be used during the assessment of their children and that may result in special education placement, parents are afforded an opportunity to voice their opinions and contribute to the final decisions when the teams (of which parents are key members) determine (a) eligibility (i.e., declare whether students meet the specific predetermined criteria for placement under a qualifying category of disability as defined in IDEA), (b) accommodations that will be necessary to meet students' individual needs (as determined by goals and objectives outlined in the students' team-generated individualized educational plans [IEPs]), and (c) placement (i.e., where children will receive their educational programming, e.g., resource room, general education classroom, some combination thereof).

Once students are identified and placed under the umbrella of special education and related services, they are entitled to a reevaluation at least every 3 years from their initial or latest evaluation to determine whether they continue to meet eligibility criteria for services. Parental permission for reevaluation is mandatory. Parents are also entitled to receive regular reports on their children's progress, and parents may request changes to educational programming (i.e., revisions of children's IEPs) to address any lack of progress toward annual goals. This level of parent recruitment and participation is significantly different from the hands-off approach popular until the late 1960s and early 1970s described earlier. As such, this legislation represents a powerful impetus for change with regard to parent and teacher roles throughout the educational process.

3.1.2. No Child Left Behind

"No Child Left Behind" (Public Law 107-110) is the latest manifestation of the public education law called the Elementary and Secondary Education Act (ESEA). As with IDEA, No Child Left Behind designates roles and responsibilities for parents and educators that necessitate home–school communication and cooperation. Perhaps the most salient feature of this legislation is its emphasis on providing information to parents while at the same time holding schools accountable for the quality of education they provide to all students. In addition to mandating that schools provide information to parents regarding their children's individual progress, No Child Left Behind requires states and school districts to give parents information about the relative strengths of each school. Specifically, for each school, they must provide achievement data (combining the performances of all students) broken

out by race, ethnicity, gender, English-language proficiency, migrant status, disability status, and income status. If, on receiving this information, parents believe that their children are not receiving adequate services, they may opt to place the children in better performing schools or tap additional resources such as tutoring and after-school programs.

Although there are clear benefits to educators and administrators (e.g., they are able to have detailed information about each child from which they can adapt teaching techniques, they are able to determine what does and does not work and make decisions based on the data), the emphasis of this legislation is on protecting individual students, parents, and families through providing critical information and options. This legislation provides critical information for educational partners to use in making decisions regarding how to improve the quality and fairness of the educational system. It also provides equal access to information, ensuring that no party has the upper hand. As indicated earlier, equal participation is a defining characteristic of educational partnerships. In addition, Goal 8 of the 1999 National Education Goals report concerns parental participation and states, "By the year 2000, every school will promote partnerships that will increase parental involvement and participation in promoting the social, emotional, and academic growth of children."

In addition to clauses mandating the types of information to which parents are entitled, No Child Left Behind includes clauses mandating that parents have the right and responsibility to be involved at the school, district, and state levels. Specifically, No Child Left Behind states that every school must have a written parent involvement policy developed with and approved by parents. If developed and executed properly, this policy affords parents much power in decision making and in determining educational policy at a local level. No Child Left Behind also calls for a school–parent compact, developed with and approved by parents, that describes how educators and parents will build a partnership to improve student achievement. If a partnership philosophy guides this compact, it can open to doors to multiple levels and types of home–school collaboration, expanding and evolving over time.

No Child Left Behind has resulted in an increased emphasis on site-based management. Site-based management, an approach with a long history, is concerned with decentralizing decision making by transferring authority and resources from state- and district-level offices to more localized units such as governance councils, committees, and teams located within each school building. This transfer of authority and power to a local level affords educational partners the opportunity to consider the data generated by schools to meet mandate requirements of No Child Left Behind and to make budget- and curriculum-related decisions at an individual school level. These site-based management teams, often composed of principals, teachers, parents, and other professionals, make decisions regarding policies for the schools. This structure necessitates input from parents as decision-making team members. Similar teams exist at the district and state board levels, whereby parents are involved as active developers and monitors regarding school board functions as well as district- and statewide governing bodies.

No Child Left Behind embodies the state of the nation regarding the pivotal role that parents play in the education of their children. This level of parental involvement requires communication and collaboration between parents and school-based personnel. As with IDEA, a complete discussion of No Child Left Behind related to home–school collaboration is beyond the scope of this article. Specific components of the legislation were discussed previously as a means of describing the rights and responsibilities of parents in education as mandated by federal law. How parents respond to such roles varies considerably across individuals and is likely influenced by multiple factors, some of which are described in the next subsection.

3.2. Home- and Family-Related Factors

There are many family-related factors that may influence how a family interacts with educators. The quality of the interaction is perhaps most heavily influenced by the family's history with the school, including how the family has been received by the school in the past. For example, during past interactions with the school, was the school hierarchical, that is, acting as if the parents' presence was an intrusion? Alternatively, was the family made to feel comfortable in the school setting? Has the family been proactively involved with the school (i.e., not just when their child is in need of assistance)? Does the family have a child who has been identified as at risk for, or as an individual with, a specific disability? If so, how was the family included in the related meetings and educational programming sequence? In general, the more positive the family's past experiences, the more likely the family is to feel accepted and welcomed as an educational partner.

Other factors that can potentially influence the nature of home–school collaboration are found under the broad umbrella of family culture. As used here, culture extends beyond race and ethnicity to include factors such as acculturation level, beliefs, values, and expectations. Acculturation level is concerned with families' adjustment to the immediate culture when they are not originally from the United States or when they are new to a particular region of the country. If they come from a place where the educational system varies considerably from that in their new surroundings, it may take them time to adjust to the new local norms regarding parental involvement in education. Some families might not view their role as that of active participants and may leave the majority of decisions and actions to educators, whereas others might take an active directive role in working with educators. Family culture also includes beliefs, values, and expectations regarding education and how to interact with the school system. Stemming from cultural, economic, and political influences, these factors are multidetermined. Regardless of whether families consist of first-time parents or parents who have been involved in the educational system long enough to witness significant changes in roles for parents in education, families will require time to adjust to their new roles.

3.3. School-Related Factors

The school is a dynamic system that is composed of multiple individuals and influenced by multiple forces (e.g., culture, economic conditions, politics). When considering school-related factors that may influence the nature of home–school collaboration, it is important to take into account both variables related to individuals within the school and building-specific traits and characteristics. With regard to school personnel, these individuals acquire beliefs, values, and expectations that are influenced by a myriad of cultural, economic, and political influences. These beliefs, values, and expectations evolve over time and influence interactions with parents. In addition, just as parents' behaviors are influenced by past experiences with the school, teachers' behaviors are influenced by past experiences with parents and families.

In addition to factors related to individuals within the school, factors related to the school itself may affect the nature of home–school collaboration. Just as each family is unique in its history and culture, each school takes on an identity that makes it unique. For example, is the school public or private? The choice to send children to a private school may represent different values and have implications for how parents and educators interact. It is also true that the size and organizational structure of the school may influence the degree to which parents feel comfortable. In general, a larger school is less personal, and this may cause parents to feel somewhat overwhelmed, resulting in less contact and collaboration with the school. The organizational structure also influences how parents interact with school personnel. For example, are teachers and administrators readily available? Are there individuals who take it on themselves to make regular contact with parents? Some school administrators may encourage teachers to maintain regular contact with parents, whereas others may value a top-down approach in which parent contact occurs through administrator-initiated actions and activities (e.g., newsletters, organized meetings). Some schools are warmer, more inclusive, and generally more family-friendly than are others. How a school "feels" to parents and students is influenced by multiple factors, and this makes pinpointing salient factors a challenge.

3.4. Community-Related Factors

Community factors that influence interactions between parents and educators are difficult to isolate given their identification across multiple systems. However, one cannot overlook the powerful influence of the community in shaping individual and systemic behaviors. Broadening the definition of cultural influences, the community system posits its own set of norms, values, and expectations regarding home–school collaboration. Behaviors, attitudes, and values transmitted from individuals in the home and school settings culminate to form a message that guides community behaviors and actions, and this in turn influences individual parents and teachers engaged in collaboration. These may be based on religious, political, or other influences and may vary over time. In addition to influencing how parents and educators interact, the community has the power to support or hinder goals and objectives identified collectively by parents and educators. For example, if parents and teachers agree that having an after-school program might help to reduce problems related to latchkey children (i.e., children who care for themselves after school), will the community embrace this plan? Does it offer supports to make the vision a reality? Do role models exist in the community to encourage more advanced collaboration? Each community has a history that guides immediate behavior, which in turn affects future interactions between various

systems and individuals. This broad-based system is pivotal and cannot be underestimated with regard to the power it has in guiding home–school interactions.

4. APPROACHES TO FOSTERING AND MAINTAINING HOME–SCHOOL COLLABORATION

4.1. Underlying Assumptions, Roles, and Responsibilities

Stemming from a myriad of cultural, economic, and political influences, the educational system has called on parents to actively participate in both the general and special education processes. Parents are increasingly becoming more actively involved in planning and decision making at the local, district, and state levels. In addition to participating in traditional activities such as parent–teacher conferences and back-to-school nights, parental involvement occurs through volunteering, planning, and serving on various committees. In addition to these responsibilities, parents of children who are identified with specific disabilities are involved in the assessment and identification of their children's specific strengths and limitations as well as in co-constructing and facilitating their children's IEPs. Parent involvement also occurs in the home and community. When parents read with their children, observe and participate in play, and support homework completion, this type of parent involvement supports the education of their children. Identifying meaningful ways in which parents can participate in their children's education is a task assigned to both parents and educators. If parents and educators engage in an educational partnership, team members will collectively identify and use specific areas of expertise that each member brings to the collaboration process as they make progress toward a mutually determined set of goals and objectives. Furthermore, if a partnership philosophy guides home–school interactions, parents and educators will seek ways in which members of the school system can support the home system.

4.2. How to Involve Families in Education

Although state and federal legislation mandates specify roles, rights, and responsibilities for parents in education, they do not prescribe specific means for establishing relationships between the home and school settings. Building a partnership is a complex process that does not follow a "cookbook" approach; rather, fostering this relationship is a process that must be approached with care based on a specific guiding philosophy. One such philosophy was described by Christenson and Sheridan, who in 2001 maintained that partnership is not an approach to home–school interactions; rather, it is a guiding philosophy for educators and parents as they attempt to build a meaningful long-term relationship. Specifically, these authors articulated four components necessary for building meaningful partnerships: approach, attitudes, atmosphere, and actions.

Approach is a means of interacting with families that embraces the values of both the home and the school in constructing the educational experiences of children. This is concerned with educators taking an approach with families that communicates their value, unique knowledge and skills, and necessity to be involved in education. When this approach is taken, it is assumed that parents will be involved in the school curriculum and that educators will support parents so that children are successful in the home setting. Attitudes are the perceptions that families and schools have of one another. This component is concerned with identifying strengths in partners, focusing on these strengths, and assuming a joint responsibility for students' development and education. Atmosphere is the overall climate set for families and educators. Finally, the actions component involves strategies for building long-term partnerships. Actions are not merely activities that include families in a passive manner; rather, they are intended to be behaviors that, over time, actively include families in the educational process. Perhaps the most salient feature of this model is its emphasis on the processes required to build relationships over time.

5. CONCLUSION

Cultural, economic, and political influences have created an educational system that values parental involvement in the education of all students. As parents and educators have adjusted to their evolving relative roles, they have engaged in home–school collaboration as a means of working together toward achieving mutually determined educational goals for all students. As parents, educators, and other professionals (e.g., school psychologists) have searched for a guiding philosophy regarding their approach to interacting with one another, many have adopted a partnership philosophy. Partnerships are long-term relationships based on the premise that coequal power and coequal participation

are prerequisite conditions for meaningful collaborative endeavors. When this philosophy is adopted, parents and educators share resources, ideas, decision-making power, and ownership for mutually determined goals regarding both the home- and school-based educational curricula. Partnerships in education allow parents to be more intricately involved in school-based educational processes as a means of enhancing home–school collaboration. In addition, schools can assist parents with the home curriculum by participating in activities such as conjoint behavioral consultation and solution-oriented family–school collaboration. Partnerships, like all relationships, take time to form and require mutual cultivation to develop and evolve. However, they can result in both meaningful and effective educational and developmental outcomes for children and yield benefits that cannot be achieved when families and schools work in isolation. With the appropriate approach, atmosphere, attitudes, and actions, home–school partnerships can create an essential condition for optimal success.

See Also the Following Articles

Consultation Processes in Schools ■ School–Community Partnerships

Further Reading

Carlson, C. I., Hickman, J., & Horton, C. B. (1992). From blame to solutions: Solution-oriented family–school consultation. In J. S. Christenson, & J. C. Conoley (Eds.), *Home–school collaboration: Enhancing children's academic and social competence* (pp. 193–213). Silver Spring, MD: National Association of School Psychologists.

Christenson, S. L., & Conoley, J. C. (Eds.). (1992). *Home–school collaboration: Enhancing children's academic and social competence*. Silver Spring, MD: National Association of School Psychologists.

Christenson, S. L., & Sheridan, S. M. (2001). *Schools and families: Creating essential connections for learning*. New York: Guilford.

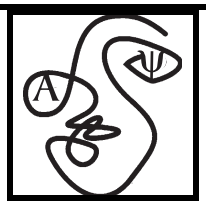
Epstein, J. L., Sanders, M. G., Simon, B. S., Salinas, K. C., Jansorn, N. R., & Van Voorhis, F. L. (2002). *School, family, and community partnerships: Your handbook for action* (2nd ed.). Thousand Oaks, CA: Corwin.

Esler, A. N., Godber, Y., & Christenson, S. L. (2002). Best practices in supporting home–school collaboration. In A. Thomas, & J. Grimes (Eds.), *Best practices in school psychology* (4th ed., pp. 389–412). Bethesda, MD: National Association of School Psychologists.

Minke, K. M. (2000). Preventing school problems and promoting school success through family–school community collaboration. In K. M. Minke, & G. C. Bear (Eds.), *Preventing school problems–Promoting school success: Strategies and programs that work* (pp. 377–420). Bethesda, MD: National Association of School Psychologists.

Telzrow, C. F., & Tankersley, M. (Eds.). (2000). *IDEA amendments of 1997: Practice guidelines for school-based teams*. Bethesda, MD: National Association of School Psychologists.

Welch, M., & Sheridan, S. M. (1995). *Educational partnerships: Serving students at risk*. Fort Worth, TX: Harcourt Brace.



Homophobia

Douglas C. Kimmel
City College, City University of New York, New York, USA

1. Heterosexist Bias in Western Culture
2. Definition of Sexual Prejudice
3. Internalized Homophobia
4. Social Attitudes, Physiological Arousal, and Homophobia
5. Gender Differences in Prejudice: Homosexuals, Bisexuals, and Transgendered Persons
6. Education, Civil Rights, and Primary Prevention of Hate Crimes
Further Reading

GLOSSARY

heterosexual bias The belief that heterosexuality is normal and that everything associated with heterosexuality is superior to anything associated with homosexuality.

homophobia Antipathy toward persons who are thought to be gay, lesbian, homosexual, or deviant from gender stereotypes in ways that suggest a same-sex sexual orientation.

internalized homophobia The negative attitude a gay man, bisexual, or lesbian holds about himself or herself as a result of growing up and living in a society with a potent heterosexual bias.

sexual minorities Lesbians, gay men, bisexuals, and transgendered persons.

sexual prejudice The negative attitude related to sexual orientation, especially toward homosexual behavior, people who have a homosexual or bisexual orientation, and the gay, lesbian, bisexual, and transgender communities.

transgendered persons Those whose gender identity, expression, or behavior is not the same as their birth sex; some experience their gender as incongruent with their anatomical sex and seek some degree of sex reassignment procedures, whereas others express their gender identity through external self-presentation and behavior.

Homophobia refers to antipathy toward persons who are thought to be gay, lesbian, homosexual, or deviant from gender stereotypes in ways that suggest a same-sex sexual orientation. It is a form of prejudice and may be related to harassment or violence toward persons thought to be homosexual. Most people in cultures that embody heterosexist values, such as Western societies, adopt attitudes that stigmatize same-sex sexual, erotic, or affectional attractions. These attitudes tend to be more prevalent in men than in women and tend to be more potent when focused on male same-sex attractions than on female ones. Frequently, these prejudicial attitudes are internalized and can lower the self-esteem of lesbians, bisexuals, and gay men. They are also associated with hate crimes toward lesbians and gay men, and in many states such crimes are considered violations of civil rights ordinances. Casual use of terms such as "faggot" may stigmatize youth in school and other settings and should not be tolerated any more than are other prejudicial and derogatory terms. Education with regard to civil rights, diversity in gender-related behavior, and tolerance of individual differences is important at all levels of school and may be most effective when it is modeled by peer leaders and authority figures.

©2004 Elsevier Inc.
All rights reserved.

1. HETEROSEXIST BIAS IN WESTERN CULTURE

Intolerance of behavior that differs from stereotypes of masculine and feminine roles is frequent in some societies but relatively absent in others. Likewise, social prohibitions against certain forms of private, consensual sexual behavior are enforced in many, but not all, cultures. The United States continues to reflect a puritanical set of attitudes about gender roles and sexual behavior. These attitudes may be seen clearly in views of same-sex erotic or affectional attractions or relationships, which have been described by some commentators as a threat to traditional social or moral values and family structure.

In the United States, persons who are perceived to have a same-sex sexual orientation can be fired from a job or denied housing, credit, or service in a restaurant or hotel unless the state or local government has an explicit law prohibiting such discrimination. Members of the U.S. armed forces will be expelled, despite their service records or other qualifications, if they reveal their homosexuality—the "don't ask, don't tell" policy. In some states, foster parents who are gay or lesbian are not allowed to adopt their foster children—or any other children—due to their homosexuality. At least one private religious university has a policy that requires students and faculty to sign a statement attesting that they are not homosexual to be allowed to study for a degree or work at the school. Likewise, other religious and secular institutions, such as the Boy Scouts of America, discriminate against homosexuals. Moreover, persons thought to be homosexual are targeted for hate crimes, including violent attacks and harassment.

In the United States during the 1960s, a growing sentiment began to emerge that homosexuals were different from heterosexuals only with regard to the sex of the persons toward whom they felt attraction and erotic love. Psychological research by Evelyn Hooker on gay men and lesbians who did not seek psychiatric treatment found no difference in psychological adjustment, and her subsequent work with the Task Force on Homosexuality of the National Institute of Mental Health began to educate professionals about the gay and lesbian subculture. At the same time, the civil rights movement, protests against the war in Vietnam, and the women's movement were challenging traditional views of social roles and institutions. Lesbians and gay men began to see the pathological view of homosexuality as something imposed on them by outside society rather than as an inherent condition of their inner selves. Likewise, they were beginning to identify as a subculture that shared characteristics with other minority groups.

During the late 1960s, a heterosexual psychotherapist, George Weinberg, coined the term "homophobia" to refer to individual anti-gay behaviors and attitudes. Herek, in an article titled "The Psychology of Sexual Prejudice," maintained that Jack Nichols and Lige Clark first used the term in print in their column on gay issues in the May 23, 1969, issue of *Screw* Magazine. One month later in New York City, a police raid on a gay bar called the Stonewall Inn ignited an extended protest against the harassment of gay people and is celebrated around the world as the touchstone of the modern gay, lesbian, bisexual, and transgender movement.

Homophobia is often criticized as being too clinical or pathological a term because it implies a pathological degree of fear (phobia). However, it might apply to a person who is afraid of being looked at or touched by a homosexual or who fears letting a child be in regular contact with a lesbian or gay male teacher. It could also involve hypervigilance or preoccupation with other people's possible homosexuality. Likewise, it might apply to the fear expressed by officers in the U.S. armed forces that having homosexuals around would damage morale because the men and women would not want to shower or be in a foxhole with a homosexual. Such irrational fears and preoccupation may indicate a true phobia about homosexuals.

The term "heterosexual bias" is frequently a more accurate term to describe the belief that heterosexuality is normal and that everything associated with heterosexuality is superior to anything associated with homosexuality. Heterosexual bias often simply ignores anyone or any situation that is not heterosexual. For example, in psychological research on dating or attraction, participants often are not even asked whether they are heterosexual because they are presumed to be so.

2. DEFINITION OF SEXUAL PREJUDICE

Recently, the term "sexual prejudice" has come to be widely adopted in research and theory. Herek, in his seminal article on the topic, defined sexual prejudice as the negative attitude related to sexual orientation, especially toward homosexual behavior, people who have a homosexual or bisexual orientation, and the gay, lesbian, bisexual, and transgender communities. He noted that sexual prejudice, as with any prejudice,

has three components in that it (a) involves evaluation or judgment (an attitude), (b) is directed at a distinct social group, and (c) involves negative feelings (e.g., dislike, hostility). Several research instruments have been designed to measure homophobia as a prejudicial attitude, using questions regarding homosexuals or lesbians and gay men included in a group of questions about attitudes toward various groups and topics.

In research studies with large random representative population samples, sexual prejudice has been found to be associated with other dogmatic attitudes, rigid views of gender roles, conservative religious views, and authoritarianism. Herek found it to be more prevalent among those persons who are older, less educated, and living in rural areas and the southern and midwestern states in the United States. Sexual prejudice was also found to be more prevalent among persons with a fundamentalist religious orientation and among those who attend religious services frequently. In some studies, persons who believe that sexual orientation is the individual's "choice" tend to be less tolerant in their attitudes toward homosexuality than do those who believe that it is beyond the individual's control. Likewise, persons who have had personal contact with lesbians or gay men tend to hold more positive attitudes than do those who have had little or no contact of this kind.

Anti-gay attitudes are, in many settings, the only socially acceptable prejudice that can be openly expressed in the media and in public statements. Fear of homosexuals is often used by right-wing organizations to boost fund-raising and is sometimes used to stigmatize a political opponent.

Society likely pays a price for the effects of sexual prejudice, including the rigid enforcement of gender roles and stereotypes, the stigmatization of long-term same-sex relationships, and the secrecy, deceit, and hypocrisy forced on those in the armed forces and other careers. In contrast, reducing sexual prejudice can enrich society by providing new models of intimate relationships and gender roles for heterosexuals as well as for bisexuals and homosexuals. It may also remove the confusion often expressed between homosexuality and gender role stereotypes such as a lack of "masculinity" or "femininity" and homosexual sexual orientations.

3. INTERNALIZED HOMOPHOBIA

A particularly interesting focus of research is on internalized homophobia, which is the negative attitude that a gay man, bisexual, or lesbian holds about himself or herself as a result of growing up and living in a society with a potent heterosexual bias. One line of research has focused on the effects of living as a sexual minority in society and the consequences that this minority stress may have on lesbian, gay, and bisexual individuals. To study this, Meyer in 1995 used an Internalized Homophobia Scale (IHS) that measures the extent to which gay men have discomfort about their sexual orientation and their efforts to avoid homosexual feelings. Two sample items included the following: "How often have you wished you weren't gay?" and "Have you thought that being gay was a personal shortcoming?" Internalized homophobia seems to be prevalent. Meyer found that 70% of a sample from the gay community expressed some internalized homophobic attitudes, although most had relatively low scores on the scale. In a study of respondents from the lesbian and gay community by Herek and colleagues, lesbians were found to have lower scores on the IHS than were gay men. For both groups, higher scores were associated with less disclosure to heterosexual friends, fewer connections with the lesbian and gay community, and higher scores on a measure of depressive symptoms and demoralization. Gay men with high scores also scored lower on a measure of self-esteem.

Internalized homophobia is a particular risk when young people first recognize their same-sex attractions. Typically, they have learned to fear and hate anything associated with homosexuality, so when they find that they might be a homosexual, they risk turning those negative attitudes in toward themselves by means such as increased incidence of depression, suicidal thoughts and attempts, and substance abuse. Lesbians, gay men, and bisexuals also are at risk for internalized homophobia if they have been victimized by hate crimes directed toward them due to their perceived sexual orientation, thinking that they have in some way deserved being victimized in this way. Although surveys have found relatively low levels of internalized homophobia in lesbians, gay men, and bisexuals over 60 years of age, it is thought that persons are at risk for internalized homophobia throughout their lives.

4. SOCIAL ATTITUDES, PHYSIOLOGICAL AROUSAL, AND HOMOPHOBIA

As with all attitudes, sexual prejudice is learned from peers, cultural media, and authority figures. It is likely

that pronounced homophobia has deeper roots than do other prejudices, however, given that it is so directly connected to gender stereotypes and sexual behavior, both of which are powerful themes in Western culture. For example, terms such as "faggot," "sissy," and "pussy," when directed at boys who are not sufficiently competitive or aggressive, may be using homophobic imagery to maintain gender role conformity. The male role is often reinforced by homophobic imagery as it frequently becomes defined by an avoidance of anything "feminine," including emotions that suggest vulnerability, tenderness, or uncertainty. Likewise, making an athletic woman who prefers short hair and is not interested in dating men to be the butt of jokes and sexual innuendos may be an example of homophobia as harassment. In both examples, homophobia is often used to reinforce gender stereotypes and to stigmatize those who deviate from the so-called traditional male and female roles. In a society that is becoming more gender irrelevant, this use of homophobia is actually an attempt to retard this change and may explain why it is stronger among both less educated and more conservative people.

Homophobia is also often linked with sexual behaviors and acts that are associated with homosexuals, especially male homosexuals. Those sexual acts, when the male is in the "passive" role, are highly stigmatized and are frequently described as ultimate humiliations of a male. Conversely, when performed by a female, even with another female, the same acts (e.g., oral sex, anal sex) may be regarded as erotic or even positively valued. Language is filled with expressions of fear regarding male anal exposure or penetration, so men try to "protect their asses" at all costs. Likewise, anything negative is said to "suck." Although the term has now mostly lost its sexual connotation, its origins in the denigration associated with oral sex are obvious.

Psychoanalytic perspectives have long suggested that persons who are most virulent in their prejudice about some stigmatized group may be projecting their own inner conflict onto that group, although they are conscious of neither their own conflict nor its projection. For example, a man who feels same-sex sexual attraction, but who cannot accept that in himself and so represses the attraction, might project his own hatred of his same-sex attraction onto someone who he perceives to be homosexual. Expression of hatred toward the homosexual would have the added benefit of deflecting any suspicion of his own same-sex feelings by others who may perceive him as anti-gay. Lesbians, bisexuals, and gay men often must work with, or be supervised by, persons such as this and are required to hide their pro-gay feelings.

One study provided empirical support for the idea that homophobic men may secretly be erotically excited by same-sex sexual acts. The experiment by researchers at the University of Georgia involved a device that measured change in the circumference of the penis and indicated the extent of erotic arousal as the men watched a series of sexually explicit videotapes. The men were also asked to rate their level of arousal by the tapes. Surprisingly, 54% of the men who scored high on a measure of homophobia before the study were definitely aroused while watching the same-sex male videos, according to the measure of penis erection, compared with 24% of the men who scored lower on homophobia. Despite their penile erections, the more homophobic men did not indicate a subjective erotic arousal by those tapes, rating their arousal as low as did the less homophobic men who showed no significant penile erections. Many of the men responded both subjectively and physiologically to videotapes of same-sex female sex acts and to male–female sex acts; these responses were unrelated to their scores on homophobia. It is possible, of course, that anxiety or other strong emotions—and not erotic stimulation—at the sight of male–male sex videos produced penile erections in the more homophobic men who did not experience subjective erotic arousal.

5. GENDER DIFFERENCES IN PREJUDICE: HOMOSEXUALS, BISEXUALS, AND TRANSGENDERED PERSONS

In general surveys of attitudes within the United States, heterosexual men express more anti-gay attitudes than do heterosexual women. Heterosexual women do not express more negative attitudes toward gay men than they do toward lesbians, whereas heterosexual men typically hold much more negative attitudes toward gay men than they do toward lesbians. A similar pattern was found in attitude research among African Americans, where the men had especially negative attitudes with regard to gay men, believing that their behavior was "unnatural." It is thought that men who violate the male gender role, which is a highly valued role in Western society, are stigmatized more than are women who violate the female gender role. Moreover, because female gender roles are less rigidly defined

than male gender roles, there may be less concern among men about lesbian women or among women about gender role violations in general.

Prejudice also exists toward bisexuals, that is, persons who feel erotic or affectional attraction to either sex. This prejudice has been termed "monosexism" to reflect the view that either homosexuality or heterosexuality is better than a combination of both orientations. Some homosexuals even believe that bisexuals should be encouraged to admit that they are actually gay, as if bisexuality were not a legitimate orientation. Herek found that attitudes toward bisexuals are more negative than those toward homosexuals. Women rated homosexuals more favorably than they rated bisexuals of either gender, and men rated males more negatively than they did females regardless of whether they were bisexual or homosexual. In general, bisexuals were rated lower on a "feeling thermometer" than were all other groups, except injecting drug users, in the survey questionnaire.

Little research has been done on attitudes toward transgendered persons, that is, those who identify themselves as neither male nor female, those who were previously male or female but have changed or are changing gender, those who have both male and female identities and wardrobes, or those whose bodies have characteristics of both sexes. A poll in 2002 found that a slim majority in the United States believed that it is "all right" to be transgender, but more than one-third believed that it is wrong. Those who believed that being transgender is a matter of the individual's choice were more likely to see it as a moral issue than were those who believed that transgendered persons were born that way. In some cultures, such persons are regarded as "two-spirit" persons and are given roles in society that fit their unique gender.

6. EDUCATION, CIVIL RIGHTS, AND PRIMARY PREVENTION OF HATE CRIMES

Research has shown that personal contact with lesbian or gay persons, such as friends, family members, classmates, and colleagues, is associated with lower degrees of sexual prejudice. Therefore, disclosing one's sexual orientation to associates can be effective education, especially if a friendship exists before or after the disclosure is made. Likewise, having homosexual peers speak in schools or social group meetings can be effective education. Finally, recognizing that most families have one or more homosexual members often helps individuals to understand that sexual prejudice may affect everyone's friendship and family circles.

Sexual minorities, including lesbians, gay men, bisexuals, and transgendered persons, are sometimes a protected class under civil rights laws. Harassment, bullying, threats, and verbal or physical violence are generally illegal and may be the grounds for civil suits. Because school districts have been held responsible for protecting sexual minority students from persecution, they are beginning to take homophobia seriously.

Hate crimes related to sexual prejudice, such as the murder of Matthew Shepard in Wyoming, have drawn attention to the most destructive effects of homophobia. Other forms of victimization related to sexual orientation can also have lasting effects and, in some states, are punished more severely than are non-hate crimes. Primary prevention efforts are needed to make communities safe for sexual minorities and to minimize the cost to society of the discrimination and prejudice that has long stigmatized lesbians, gay men, bisexuals, and transgendered persons.

See Also the Following Articles

Homosexuality ■ Risk-Taking and Sexual Behavior ■ Sexual Behavior ■ Sexual Behavior and Culture

Further Reading

Adams, H. E., Wright, L. W., Jr., & Lohr, B. A. (1996). Is homophobia associated with homosexual arousal? *Journal of Abnormal Psychology*, *105*, 440–445.

Garnets, L. D., Herek, G. M., & Levy, B. (1990). Violence and victimization of lesbians and gay men: Mental health consequences. *Journal of Interpersonal Violence*, *5*, 366–383.

Garnets, L. D., & Kimmel, D. C. (Eds.). (2003). *Psychological perspectives on lesbian, gay, and bisexual experiences* (2nd ed.). New York: Columbia University Press.

Greene, B. (2000). Beyond heterosexism and across the cultural divide: Developing an inclusive lesbian, gay, and bisexual psychology—A look to the future. In B. Greene, & G. L. Croom (Eds.), *Education, research, and practice in lesbian, gay, bisexual, and transgendered psychology* (pp. 1–45). Thousand Oaks, CA: Sage.

Herek, G. M. (2000). The psychology of sexual prejudice. *Current Directions in Psychological Science*, *9*(1), 19–22.

Herek, G. M. (2002). Heterosexuals' attitudes toward bisexual men and women in the United States. *Journal of Sex Research*, *39*, 264–274.

Herek, G. M., & Capitanio, J. P. (1996). "Some of my best friends": Intergroup contact, concealable stigma, and

heterosexuals' attitudes toward gay men and lesbians. *Personality and Social Psychology Bulletin*, *22*, 412–424.

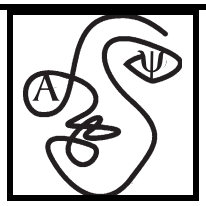
Herek, G. M., Kimmel, D. C., Amaro, H., & Melton, G. B. (1991). Avoiding heterosexist bias in psychological research. *American Psychologist*, *46*, 957–963.

Kite, M. E., & Whitley, B. E., Jr. (1998). Do heterosexual women and men differ in their attitudes toward homosexuality? A conceptual and methodological analysis. In G. M. Herek (Ed.), *Stigma and sexual orientation: Understanding prejudice against lesbians, gay men, and bisexuals* (pp. 39–61). Thousand Oaks, CA: Sage.

Meyer, I. (1995). Minority stress and the mental health in gay men. *Journal of Health and Social Behavior*, *7*, 9–25.

Rothblum, E. D. (2000). "Somewhere in Des Moines or San Antonio": Historical perspectives on lesbian, gay, and bisexual mental health. In R. M. Perez, K. A. DeBord, & K. J. Bieschke (Eds.), *Handbook of counseling and psychotherapy with lesbian, gay, and bisexual clients* (pp. 57–79). Washington, DC: American Psychological Association.

Savin-Williams, R. C. (1999). Matthew Shepard's death: A professional awakening. *Applied Developmental Science*, *3*, 150–154.



Homosexuality

Bertram J. Cohler and Phillip L. Hammack

University of Chicago, Chicago, Illinois, USA

1. Introduction
2. History, Culture, and the Problem of Defining Homosexuality
3. Understanding Same-Sex Desire: The Search for Biological Origins
4. Realization of a Gay or Lesbian Identity and the Life Course
5. Childhood and Adolescence
6. Homosexuality and Society
7. Conclusion

Further Reading

GLOSSARY

coming out Process of self-acknowledgment and disclosure to others, including friends and family, of attraction to, and desire for a sociosexual relationship with, another person of the same sex.

gender-stereotyped interests Childhood play and school and sports activities presumed to be appropriate for own gender; they are most often characterized as "rough-and-tumble" play for boys and as doll play and quiet activities for girls.

sex role Recognition and enactment of the socially constructed set of behaviors traditionally considered appropriate to one's own gender.

sexual identity Construction founded on the confluence of social and historical circumstances and personal experiences of a sense at any point across the course of life of membership in that group of persons sometimes called homosexual, gay, lesbian, queer, or spectrum or in that group of persons sometimes called heterosexual or straight.

sexual lifeway Constellation of culturally defined expectations regarding the life course for persons self-identifying with a particular sexual orientation such as gay, bisexual, or lesbian.

sexual orientation Enduring awareness of sexual attraction to, or desire for, others of the same or opposite sex.

Homosexuality refers to both a socioerotic attraction to others of the same gender and an identity as a member of a community sharing this same-sex socioerotic attraction. Although it is a term preferred by men and women in earlier generation cohorts, younger cohorts view this as a medical term and one that implies opprobrium. Members of cohorts born during the postwar era and coming of age from the 1970s to the present prefer a term such as "gay" or "lesbians and gay men." Members of the youngest cohort, born during the 1980s, often prefer a term such as "queer" or "spectrum," referring to an other than normative (heterosexual) identity and beyond simple classification.

1. INTRODUCTION

The term "homosexuality" refers to the experience of exclusive or nearly exclusive erotic preference for others of the same sex in fantasy and, characteristically, through realization of sexual intimacy with others of the same sex. It can be conceptualized as desire, behavior, and identity (although it is not always congruent

©2004 Elsevier Inc.
All rights reserved.

within one person at any one time and so may represent a source of personal conflict). Homosexual desire is the affective experience of same-sex attraction. Homosexual behavior occurs when members of the same sex engage in sexual activity. Homosexual identity represents the assumption of a self-label in which one acknowledges the primacy of homosexual desire and behavior in his or her self-composition. In contrast to homosexual identity, homosexual orientation reflects the affective experience of primarily same-sex erotic desire beyond conscious control and outside the purview of historical and social construction.

The experience of same-sex sexual contact has occurred across cultures and historical eras. As such, it is difficult to generalize the meaning of homosexuality from one society or one time to another society or time. Even in American society, the meaning of homosexuality has undergone radical transformation and permutations during the past century. The current discussion is focused largely on the experience of men and women in contemporary Western society who self-identify as "homosexual" or, more recently, as "gay," "lesbian," or "queer." This identity assumption in postmodern Western culture represents a particular sexual "lifeway" bound by culture and history. Thus, one's conception is necessarily limited and rooted primarily in the experience of urban bourgeois gay men and lesbians in North America and Europe because these are the voices represented nearly exclusively in the literature on homosexuality. This article begins with a historical account of homosexuality in the current cultural context, followed by a discussion of the controversies regarding the origin of homosexuality. Next, it reviews theory and research on the gay and lesbian life course, beginning with the "coming out" process. Finally, the article considers four contemporary "problems" pertaining to gay men and lesbians related to the larger social and political ethos: mental health, social stigma and hate crimes, acquired immune deficiency syndrome (AIDS), and marriage.

2. HISTORY, CULTURE, AND THE PROBLEM OF DEFINING HOMOSEXUALITY

In much of the public health and gay and lesbian psychological literature, the emergence of the phrase "men or women who have sex with other men or women" has emerged in response to the phenomenon of consistent homosexual behavior in the absence of homosexual self-identification. For example, many young adults eschew any sexual identity as a "natural kind," represented in polarizing labels such as "gay" and "straight." The term "queer," as used by these young adults, is meant to defy such essentializing efforts and reflects a new view of sexual desire as less amenable to simplistic bifurcations.

Recent research and theory have highlighted the importance of studying gay and lesbian lives within the context of shared, cohort-specific understandings of sexual identities and lifeways. Following a series of more gradual shifts in the social organization of sexuality during the first half of the 20th century, gay and lesbian life exploded into the public consciousness following the advent of the gay rights movement during the late 1960s and early 1970s. Inspired by and modeled on the civil rights movement, gay liberation was ignited in the aftermath of a 1969 riot in a New York City bar, the Stonewall Inn. Although the gay rights movement, symbolized by the Stonewall Inn riot, forever altered the shape of gay and lesbian life, its impact was largely cohort specific. Not only are different generations defined by different historical moments, but they also react in somewhat distinct ways to the same historical events. Over the subsequent three decades or so, there was a dramatic change in discourse regarding sexuality and sexual identity that has led to recognition of same-sex socioerotic identity as "virtually normal."

One's understanding of homosexuality today is highly influenced by the cultural and historical context of the late modern West. In particular, World War II created a context in which individuals with homosexual desires who perhaps felt alone in their experience were able to discover one another. Following the war, they settled in large metropolitan centers and began to form distinct gay and lesbian communities or subcultures. The emergence of a distinct homosexual identity as a clearly definable social category, as well as the growth of the gay and lesbian community, was fueled in large part by the social ethos of the 1960s and its culture of sexual liberation. The resistance of larger social institutions to accepting same-sex intimacy or homosexuality as a legal, legitimate social identity inevitably led to the Stonewall Inn riot, which sparked the beginning of the gay civil rights movement. Within the backdrop of this cultural and historical context, individual lives unfolded, and tantamount to social acceptance of gay men and lesbians has been the search for a causal mechanism—one that would ultimately demonstrate the immutability of homosexual desire.

3. UNDERSTANDING SAME-SEX DESIRE: THE SEARCH FOR BIOLOGICAL ORIGINS

Our culture is one in which lives, like stories, are assumed to have a linear organization with beginnings, middles, and ends—or childhood, adolescence, adulthood, and old age. Our society is one in which particular emphasis is placed on beginnings and origins. From learning theory (with an emphasis on primacy and recency of learning, ethology, and the concept of the critical period) and psychoanalysis (with a stress on the nature of the experience of early ties with caregivers), personality theory and developmental study has focused on origins and "causes" of particular personality characteristics. Nowhere is this concern with origins more evident than in the search for the causes of same-sex desire. This concern has been manifest both in the study of presumed biological determinants of homosexuality and in the discussion of the early development of the boy or girl later destined to choose same-sex sexual partners.

The debate in research and theory on homosexuality between the essentialist and constructionist positions is particularly concerned with the question of origins. The more politically controversial constructionist position emerges from social theories of identity as created through practice, suggesting that gay identity is formed within and among individuals as a consequence of shared cultural practice. The essentialist position, in contrast, posits that homosexuality is an innate condition of human experience. The biological line of research on homosexuality is representative of this position. Because of its controversial nature, including the implication of gay men and lesbians as "others" with some congenital defect (believed by some gay activists as fostering the cause of gay and lesbian political rights), the biological account of homosexuality is explored in depth in this section.

3.1. The Four Areas of Biological Study of Homosexuality

The search for biological determinants of same-sex desire began during the 19th century and has continued to the current time with reports in scientific journals. Interest in this has been enhanced by those who believe that if a biological determinant can be found, same-sex desire might be viewed as an inherent and unchangeable attribute, much like hand preferences (right or left) or other differences, and then subject to legal protection. Significantly, although there have been scattered references in the literature to the biological determinants of homosexuality among women, nearly the entire corpus of biological study focuses on the determinants of homosexuality among men.

The four areas in which biological contributions to the study of the origins of homosexuality have been reported are social biology and evolutionary theory, behavioral genetics and the search for the "gay" gene, hormones and prenatal development, and brain structure and anomalies from "normal" brain anatomy. Sociobiological models maintain that because the gene for homosexuality has not disappeared from the population, this gene must have some adaptive value. Claiming that gay men are more potent than their heterosexual counterparts, then, a heterosexual man who is heterozygous for the gay gene would have a reproductive advantage over a straight counterpart who is not carrying the gay gene. Other sociobiological theories suggest that parents recognize that some offspring are less biologically fit (e.g., lower intelligence, lower social adeptness) than others and so protect these children from challenging life experiences such as parenthood and "make" them into homosexuals. This perspective is virtually silent regarding homosexuality among women.

A second biological model focuses on behavioral genetics. Building on the observation that there is a "family disposition" toward homosexual interests, early reports contrasting monozygotic (MZ) and dizygotic (DZ) twins claimed that means of genetic transmission could be understood from a Mendelian perspective. More recent behavioral genetic studies have assumed polygenetic transmission. In 1991, Bailey and Pillard reported that the rate of concordance of homosexuality among MZ twin pairs is approximately 50%, whereas the rate among biological siblings is 22% and that among adopted brothers is 11%. This research group has also reported concordance rates for homosexuality among women who are MZ and DZ twins approximating those for homosexual men. After reports showing enhanced concordance for homosexuality among MZ twins as contrasted with DZ twins or brothers, Hamer and colleagues in 1993 reported a study of possible genetic linkage following the logic that if a trait is inherited through a single gene, relatives who share the trait will share the gene more often than would be expected by chance. Even if a gene that is responsible for a trait may not be identified, it may be shown to exist by means of transmission of another clearly inherited quality to which the trait is linked. Linkage studies are a common means for the study of inheritance. Following

this logic, Hamer and colleagues studied 40 families, each of which included two gay sons. Focusing on a gene located on the X chromosome that is transmitted from the mother but expressed only in men, the investigators located a marker on the Xq28 region that was concordant for 33 of the 40 pairs of gay brothers. In a large group of presumably straight brothers, the marker was randomly distributed. The chance of this association was less than 1 in 10,000. Hamer and colleagues maintained that this trait of homosexuality was transmitted through the maternal line, with the gene for homosexuality being located on the X chromosome, and confirmed this prediction in the reports by a group of gay men who attributed homosexuality to a greater number of maternal relatives than paternal ones.

A third area of biological study regarding the determinants of homosexuality focuses on the role of hormones, particularly during prenatal sexual development. This tradition presumes that same-sex sexual orientation, particularly among men, is a result of biosocial stress during critical periods of fetal development. Using animal models, prior observation had shown that in the absence of fetal androgens, the external genitalia develop as female, leading to the presumption that masculinization of brain centers also required the presence of these hormones. Particularly important for the development of homosexuality is the development of sex hormone receptors outside of the genital apparatus during the fourth through seventh months of gestation. Appropriate levels of testosterone appear to be necessary for these receptors to develop. The possible impact of disturbed intrauterine testosterone levels, such as might result from unusual stress or other environmental factors, has been a subject of continuing study. Aberrations in the hormonal receptor sites that might be caused by such changes would not be manifest until the time these receptors ordinarily begin to function. For example, males who had partially failed to develop testosterone receptor sites might not show the changes in muscle mass and strength that typically accompany pubertal development. Testosterone and its metabolite dihydrotestosterone (DHT) are necessary for male sexual arousal, so abnormal hormonal receptors might also influence sexual arousal. The impact of the failure to produce testosterone on the development of masculinity has been demonstrated in laboratory studies using rats.

This area of study has played an important role in work summarized in 2002 by Friedman and Downey regarding the emergence of homosexuality in men. Presuming androgen insufficiency that leads to incomplete masculinization in boys destined to become homosexual, Friedman and Downey maintained both that parents implicitly recognize these boys' feminine nature and react to them as if they were less than masculine and that these boys avoid the "rough-and-tumble" play that is characteristic of typical boys. In fact, Friedman and Downey conflated gender-atypical interests (e.g., in theater or the arts) with femininity and assumed that these boys have some elements of a gender identity disorder that later becomes manifest as same-sex sexual orientation.

A fourth area of study regarding the biological determinants of sexual orientation, primarily homosexuality among men, was stimulated by reports of differences in the first through third interstitial nuclei of the anterior hypothalamus observed on contrasting groups of gay and presumably heterosexual men. These reports suggested that the hypothalamus, long believed to be important in determining sexuality, is sexually dimorphic. This study is closely related to that regarding intrauterine androgen exposure given evidence that the size of anterior nuclei of the hypothalamus in rats depends on prenatal androgens during a critical period of prenatal development, that this area is critical for rat sexual functioning, and that reported differences in the size of these nuclei in the human anterior hypothalamus are consistent with the prenatal stress–androgen hypothesis. This research presumes that gay men must be like women in preferring men as sexual partners because, for example, the third anterior nucleus of the hypothalamus has been shown to be larger in women than in men and so should be larger among gay men, who are like women in their preference for men as sexual partners, than among heterosexual men.

3.2. Problems Posed by Biological Study of the Origins of Homosexuality

Much of the study attempting to show some biological determinants of homosexuality has reflected serious problems related both to the method of study and to the measures used to evaluate sexual orientation. For example, the social evolutionary perspective has been challenged on the ground that this perspective is mired in stereotypes and unwarranted assumptions such as that gay men are more potent and more successful in finding sexual partners than are heterosexual men. In 1999, Stein observed that in behavioral genetics study, there is an "ascertainment bias," possibly leading gay

men with identical twins both to be more likely to volunteer for such studies than gay men with fraternal twins and to overreport concordance for same-sex desire. Stein noted that concordance rates for sexual orientation are considerably lower than had been claimed when issues of volunteering were explored and that they have been inflated by sampling bias. Furthermore, whereas the genetic hypothesis predicts equal rates of homosexuality among DZ twins and brothers, Bailey's research group reported that unrelated adopted siblings were more likely than brothers to report same-gender sexual orientation (11 vs 5%). Finally, because twins were reared together in virtually all studies, it is difficult to disentangle environmental and biological effects. In sum, various problems of method, including the manner in which participants were located, reliance on indirect reports regarding sexual desire, focus on behavior rather than fantasy, measurement of zygosity by self-report, use of possibly unreliable scales for measuring sexual orientation, and very small sample sizes, all compromise the findings from this research.

Further problems are found in studies of the gay gene, including sampling methods, problems in replicating earlier studies, and problems of volunteer bias that are further compounded by the use of mail-in questionnaires and the reports of a variety of sexual behaviors among these volunteering participants. An approach to the study of the biology of homosexuality fails to consider the complex interplay of genetics, hormones, and social context and also blurs the distinction between association and causation. Reviewing the research in the area of hormones and sexual orientation, Stein concluded that there is little evidence showing even an association between expressed sexual identity and presumed hormonal factors. All of these problems are replicated anew in studies of neuroanatomy and sexual orientation.

Stein was particularly critical of the assumption made in much research on sexual orientation that this orientation is of a "natural kind" and that gay men can be readily classified as either "gay" or "straight." Stein systematically examined the assumptions underlying claims that sexual orientation is founded on some essential (biological) attributes. He observed that the dilemma inherent in understanding sexual orientation from this essentialist position is founded in the very presumption of categories (either binary or bipolar) of "natural kinds" of human sexuality, such as those of homosexual or heterosexual, man or woman, and sexual identity or gender identity, rather than recognizing the many ways in which persons understand or make meaning of their sexuality and the limitations imposed by historically and culturally available constructs to name one's own experience. There may be little merit in trying to fit people into these binary categories. As Chodorow observed in 1992, it is not even understood why men and women become heterosexual.

The politics of gender and sexuality in contemporary society have further contributed to the publication of studies believed to show that sexual categories are a natural kind. The question is not whether biology is involved (biology is involved in all aspects of human development) but rather how biology is involved in particular kinds of sexual desires according to a permissive model in which genetic factors influence aspects of neuroanatomy, which then influence particular preferences; a direct model that constrains or shapes a particular preference; or an indirect model in which genetic factors influence personality, which then interacts with life circumstances to shape sexual orientation. Stein suggested that most findings to date support an indirect model of determination. Stein questioned the assumption that persons who have sex with others of the same gender uniquely possess attributes that need to be explained or that sexual orientation may be regarded as either binary (gay or heterosexual) or bipolar (a continuum marked at one end by gay and at the other end by heterosexual).

4. REALIZATION OF A GAY OR LESBIAN IDENTITY AND THE LIFE COURSE

The process of coming to self-identify as a homosexual (or, more recently, as a gay or queer) man or woman is often termed "coming out" (of the closet). In 1994, Chauncey observed that this term may be traced back to the 1930s when it was used as a satire for the debutante experience. However, during the postwar period, this term was further refined to mean coming out into the gay life or the gay world and out of hiding (portrayed by the often used term "closet"), including first discovery of same-gender sexual experiences and, more recently, disclosure of same-gender sexual desire to family and community. The terms "gay," "lesbian," and "queer" that are preferred by many younger urban-living men and women experiencing same-sex desire reflect an effort to turn a spoiled identity into a positive one.

Coming out represents more than just the disclosure of one's same-sex desire to others. It is a process of

internal experience in which an individual's entire sense of identity is challenged, reformulated, and crystallized. Either concurrent with or proceeding the process of coming out to self, individuals engage in a cultural resocialization process, usually by attending a coming out group at a community agency or gay social services center, that results in the eventual assumption of a gay, lesbian, or queer identity. Models of the coming out process have typically assumed a developmental character, best exemplified by the stage models, and have been challenged by perspectives that are more reflexive, discursive, and social constructionist. The historical and cultural relativity of the experience of those with same-sex erotic desire questions the value of stage-based developmental models of the coming out process. Stage-based models are concerned with ideal types and ignore the complexity of intersections with other aspects of identity, such as ethnicity, and so should be used only as a very basic guide to conceptualizing the identity development process for gay men and lesbians.

Adopting a symbolic interactionist perspective from sociology, humans behave in accordance with their own perceptions of meaning that are constructed through social interaction. The idea of homosexuality represents a socially constructed category or dimension of sexual experience that resonates with individuals constantly engaged in meaning-making (i.e., symbolic) human interaction. One becomes "homosexual" in this framework through the enactment of a particular role in social interaction. Being homosexual occurs in and through practice. It is not an "essence" located within a particular anatomical or genetic substrate. Rather, it is a meaningful classification of identity formulated, negotiated, and maintained through real human social activity. The social constructionist account of sexual identity has also been promulgated by a number of scholars working with lesbians. Much of this work reveals that, for White middle-aged lesbians and bisexual women of European descent, coming out is a nonlinear process that cannot be mapped onto specific stages.

The psychological and sociological approaches are nicely complemented by anthropological perspectives on gay and lesbian identity development such as that proposed in 1997 by Herdt, who posited that coming out represents a ritual that marks the beginning of a cultural resocialization process in which individuals with same-sex erotic desire learn the history, lore, and customs of a larger gay culture. In sum, coming out represents a process that occurs both within the individual and his or her relationship to others through social interaction and within the larger gay and lesbian culture. As such, it is useful to consider the gay and lesbian identity development process from psychological, sociological, and anthropological perspectives given that they complement one another in their emphases on the individual, society, and culture.

5. CHILDHOOD AND ADOLESCENCE

Research on the childhoods of gay men and lesbians has been based largely on adolescent or adult retrospective reports. A theme that is particularly salient during the childhoods of gay men and lesbians is that of difference. As children, gay men and lesbians often report feelings of being different from peers, although they often do not define this difference in terms of sexual desire. Yet most gay men report very early awareness of same-sex attraction, typically during the elementary school years. Research on childhood experiences of homosexuality are obviously biased by this reliance on retrospective reports, but one can conclude that the affective response to same-sex desire begins during childhood.

Many investigations of lesbian, gay, bisexual, and transgendered (LGBT) adolescents have focused on issues relating to mental health and psychological adjustment. Although some studies of the mental health of gay and lesbian adolescents have suggested that, in contrast to their heterosexual counterparts, LGBT adolescents report greater levels of depression, hopelessness, and suicidal preoccupations, after controlling for environmental factors such as stress and social support, significant differences between the groups disappeared. These findings suggested that changes in resources for LGBT youth have the potential to diminish negative psychological outcomes. Studies have suggested that reactions of parents and perceptions of parental support are particularly important in predicting levels of self-esteem among LGBT youth and adolescents (as well as among LGBT individuals as whole) who are at increased risk for victimization and harassment (i.e., acts of verbal and physical violence).

Acts of victimization and harassment have been associated with a number of problems commonly experienced by LGBT youth, including school-related problems, suicide, running away, substance abuse, prostitution, and conflict with the law (usually for substance abuse, prostitution, running away, and/or truancy). When acknowledging the relationship between these problems and victimization and harassment, it is clear that LGBT youth need support systems. Organizations committed to aiding LGBT youth must also commit

themselves to breaking down the stereotypes and prejudice that result in victimization and harassment.

5.1. Adulthood and Aging

Little research has focused explicitly on the lives of lesbians and gay men in their 20s and 30s, due in part to the rapidly changing life course for gay men and lesbians. For example, with ages of coming out decreasing, gay men and lesbians are more closely mirroring their heterosexual counterparts in terms of romantic relationships. Once youth come out during adolescence, they are free to explore intimate relationships with same-sex partners by late adolescence. Early adulthood, then, is often characterized by serial monogamous relationships among gay men and lesbians, just as it is among their heterosexual counterparts.

Most discussions of adult development occur within a framework that tends to assume heterosexuality. Work is a major issue for lesbians and gay men at middle age and is also an area that differentiates them from their heterosexual counterparts. There tends to be more continuity in the professional lives of lesbians than in those of their heterosexual counterparts because lesbians' careers are less often interrupted by child rearing. However, Patterson noted in 2000 that this is likely changing as more gay men and lesbians are raising children. Most lesbians are in committed relationships and working during midlife. However, there is enormous diversity in experience during midlife among both lesbians and gay men, and little systematic research has addressed the question of midlife development among gay men.

At middle age in our society, people begin to look backward in evaluating their own lives rather than looking to the future. The "crisis" that is so often portrayed for men in our society as they reach middle age might be exacerbated for gay men as a result of their perceived loss of physical attractiveness. Sense of a foreshortened life course is common among the cohort of now middle-aged men who are survivors of the AIDS pandemic that peaked during the 1980s and early to mid-1990s and who suffered the loss of so many lovers and friends to AIDS. However, because the developmental trajectory of gay men across the adult years does not necessarily follow that of their heterosexual counterparts, it is difficult to generalize from heterosexual to gay experience of self and others across the second half of life.

The experience of later life (i.e., after 65 years of age) for lesbians and gay men is becoming increasingly investigated as the field of gerontology and aging expands exponentially (in conjunction with increased longevity). As with gay men and lesbians at middle age, the lives of older gay men and lesbians cannot be extracted from historical context and the view of the larger culture toward homosexuality. It is possible that the unique experience of coming out (a major life transition in which one evaluates the self), together with lifelong experience in dealing with stigma and anti-gay prejudice, leads to crisis competence that fosters the development of personal strengths and, among older gay men and lesbians, further leads to successful adjustment to the challenges posed in our society for older adults. Research with older lesbians and gay men reveals highly successful coping mechanisms in the face of considerable stressors and the expected course of aging.

6. HOMOSEXUALITY AND SOCIETY

This final section reviews four contemporary issues involving homosexuality and society. First, the research on homosexuality and mental health is reviewed, detailing briefly the current controversy over "conversion" therapy for gay men and lesbians. Second, the impact of social stigma on personality development is reviewed, with specific attention given to the social problem of hate or bias crime. Third, the relationship between homosexuality and the AIDS pandemic is discussed. Finally, the issue of gay and lesbian marriage, and its implications for the changing history of the gay and lesbian life course, is discussed.

6.1. Mental Health

The question of the mental health of lesbians and gay men has been a major focus of inquiry since the elimination of homosexuality as a diagnosable mental illness in 1975. As members of a social group that has been marginalized in contemporary society, gay men and lesbians continue to be at risk for a host of mental health and adjustment difficulties, especially those associated with the coming out process. For example, mental health problems that adolescents experience concurrent with coming out are experienced intrapsychically but are rooted in external sources of stress (e.g., prejudice, victimization) and lack of social support. As attitudes toward homosexuality become increasingly positive, and as institutional support for young gays and lesbians grows (e.g., through gay–straight alliances in high schools), sources of external stress will likely diminish, improving the overall coming out experience. Mental

health problems that are typically experienced during the coming out process include low self-esteem, depression, anxiety, and suicidal ideation.

A major mental health concern for lesbians and gay men involves the experience initially portrayed by Malyon in 1982 as internalized homophobia. Internalized homophobia occurs when a gay man or lesbian possesses negative feelings of self based on the larger social phenomenon of homophobia or anti-gay prejudice. That is, the individual fails to reject the anti-gay attitudes of the dominant ideology and, instead, assumes the veracity of these attitudes. Research has suggested that gay men and lesbians with higher levels of internalized homophobia are more likely to engage in self-destructive behaviors, such as substance abuse, and to experience a whole host of other adjustment difficulties. The experience of internalized homophobia is likely implicated in a host of mental health and adjustment difficulties for gay men and lesbians. In sum, the mental health of the vast majority of gay men and lesbians is comparable to that of heterosexuals. The mental health problems most commonly experienced by gay and lesbian patients center on their struggles to adapt to a heterosexist society.

Although the mental health field has a history of pathologizing homosexuality, the ideology underlying more recent clinical training and practice has been largely affirmative of gay and lesbian identity. Advocates of so-called conversion or reparative therapy have claimed success in changing the behavioral sexual orientation of gay men expressing dissatisfaction with their gay identity, although they have not been able to show that these men have given up their fantasies about same-sex experiences. Drescher's careful review of this controversial therapy in 1998, as well as the report of the Group for the Advancement of Psychiatry in 2000, highlights the problems with the assumptions on which this therapy is founded and also the validity of claims regarding presumed successful outcomes of this conversion psychotherapy. Currently, most of the major professional mental health associations regard practice of reparative therapy as problematic or even unethical.

6.2. Social Stigma and Development

A second major issue involving homosexuality and contemporary society is the impact of social stigma on personality, identity, and overall mental health. During the recent history of Western society, attitudes toward lesbians and gay men have largely been characterized by heterosexism and homophobia. In 1993, Herek portrayed heterosexism as an ideology that denigrates and stigmatizes nonheterosexual lifestyles and privileges heterosexuality. Prizing heterosex reflects a value assumption that it is the only appropriate manifestation of love and sexuality and devalues all that is not heterosexual. Herek reported that heterosexist and anti-gay attitudes appear to be more common among heterosexuals who are more religious, adhere to a more "traditional" ideology of family and gender, and have friends with anti-gay attitudes. Heterosexuals who report more contact with gay men and lesbians tend to endorse fewer anti-gay attitudes than do heterosexuals who have not had extensive contact with gay men and lesbians.

The large-scale stigmatization of lesbians and gay men exerts negative social, economic, and emotional effects on their development. At both the cultural and individual levels, heterosexism is implicated in the increased experience of victimization and anti-gay violence. Gay men and lesbians, especially youth, continue to experience victimization based on their sexual identity. The highly publicized, tragic murder of Matthew Shepard in Wyoming was a poignant reminder of the deleterious effects of cultural heterosexism and individual anti-gay attitudes. Heterosexism assumes a powerful role in both the internal (e.g., feelings of difference) and external (e.g., victimization) aspects of the lives of gay men and lesbians. The elimination of attitudes associated with heterosexism and homophobia has implications for enhancing the optimal development of lesbians and gay men by creating a context in which prejudice and the devaluing of a homosexual identity are minimized.

6.3. Homosexuality and the AIDS Pandemic

Too often, homosexuality among men has been conflated with being positive for the virus leading to outbreak of symptoms of AIDS, to such an extent that young men coming of age and experiencing same-gender sexual desire during the subsequent decades became terrified that their attraction meant certain infection and death. The virus responsible for AIDS was isolated and a test for presence of the antibodies responsible for the human immunodeficiency virus (HIV) infection became widely available by 1985. However, with no known treatment, there was some reluctance to get tested. Even after the development of the HIV antibodies test, as men began getting sick and dying, few understood

that unprotected anal sex, which facilitated blood–semen contact, was the primary means of transmission.

It is well known that AIDS is ever less a "gay" disease as many millions of persons around the world have become infected through needle use and other means of transmission but cannot afford the expensive medication available to affluent men and women in the United States. Much of the effort in the mental health field has focused on means of risk reduction through education promoting safer sex. Belief in the efficacy of safer sex leads to safer sex practices. Men who are more comfortable with their sexual orientation as gay, and who are out in the community with a lessened focus on anti-gay prejudice as a stress for them, are also more likely to practice safer sex and, in turn, receive a sense of enhanced personal mastery from such practices.

6.4. "Normal" and Beyond: Gay and Lesbian Relationships, Marriage, and Parenthood

The gay identity is often a contested one on the part of gay and lesbian people and the larger community alike. On the one hand, findings from recent surveys have shown a decrease in anti-gay bias. On the other hand, efforts within the gay and lesbian community to "fit in" within the larger society have been contested both by the larger society and within the gay community itself. Much of this controversy is apparent in the recent heated debate with the Episcopal church regarding the election of "out" gay priest Gene Robinson as a bishop of the church. However, this contested identity as gay or lesbian is most clearly evident in the area of same-sex partnership. Marriage in the United States symbolizes both social order and compliance with traditional norms. Even though divorce marks the end of one in three marriages, shared beliefs reflect the larger belief that marriage has a special meaning in terms of prevailing concepts of love and romance between husband and wife.

Given the significance of this belief in the special quality of the marital tie, there was wide-scale protest in 1993 when the Hawaii Supreme Court held that same-sex couples shared the right with heterosexual couples to marry under the equal protection clause of the state's constitution. This decision prompted a firestorm as many states and the federal government passed legislation banning same-sex marriages. In 1996, Congress passed the Defense of Marriage Act, signed into law by President Bill Clinton, that denied legal recognition to same-sex unions at the federal level. The Catholic Church and most mainline Protestant churches prohibit clergy from performing marriage rites with same-sex couples. At the same time, there is increasing support for legal recognition of same-sex civil unions beyond support of domestic partnership, as was recently approved in Vermont. It appears that the word "marriage" has some symbolic significance reflecting traditional norms ever more important when confronting the reality of social change.

In November 2003, the Massachusetts Supreme Court held that a prohibition on same-sex marriages violated the equal protection clause of the state's constitution. The court upheld the equality of all persons and noted that marriage confers legal, financial, and social benefits, and also brings stability, to society. The court saw no constitutional reason why such benefits should be denied on the basis of sexual orientation. After the Vermont decision to legally recognize same-sex civil unions came a decision by Canada's province of Ontario to recognize same-sex marriages (soon to be extended across the country), legal recognition of same-sex marriages in The Netherlands, a move for such legal recognition in Australia, and a move for legal protection for same-sex unions in Scandinavia, there is emerging support for legal recognition of gay marriage. In response to this trend, federal officials in the United States, from the president to a majority of Congress, continue to be opposed to same-sex marriages, and the Massachusetts decision has sparked renewed discussion regarding a U.S. constitutional amendment that would define marriage as a union between a man and a woman.

Posner, a federal appeals court judge and legal scholar, maintained in 1992 that government must not discriminate against classes of citizens. Although he appeared to be in favor of same-sex marriages, he indicated that society was not yet ready for this change and suggested granting legal protection to same-sex couples, such as in granted in Scandinavia, as an interim solution. Underlying these legal decisions there is continuing controversy even within the gay community regarding same-sex marriages. In 1997, these disparate positions were included in Sullivan's edited collection that includes summaries of important court cases, congressional debate on the Defense of Marriage Act, and passages from the works of both supporters and critics of same-sex marriages. In 1995, Sullivan himself argued in favor of same-sex marriages, noting that extending legal protection would enhance the stability of same-sex relationships. He refuted the position of many opponents to same-sex marriage that

homosexual lifeways are necessarily depraved and the argument that remaining single is a better social outcome for gay men and lesbians than is entering into a legally recognized union that would enhance personal and social stability. In 1999, Warner indicated a concern that such legal recognition might undermine the distinctive culture created by gay men and lesbians and noted that the high divorce rate undermines the argument that heterosexual marriage fosters social stability.

The court in the Massachusetts decision explicitly commented that a ban on same-sex marriages proved to be a disadvantage for children of same-sex unions. Indeed, one area in which legal protection for gay and lesbian unions would be particularly important concerns parenthood and child care. One of the numerous areas in which the social changes of the past decade have been significant concerns birth, custody, and adoption among gay and lesbian couples. Custody decisions in divorce cases have generally favored the wife and mother, even when she has self-defined as lesbian and has sought divorce so as to join her lover. In 2000, a report of the Group for the Advancement of Psychiatry on the problem of bias and mental health among gay men and lesbians reported that several states have ruled that sexual orientation does not by itself imply unfitness of a parent in a contested custody proceeding and that parental unfitness and evidence of harm to the child must be demonstrated. In the past, courts have been reluctant to award custody to the father in divorce cases. Most often, where gay couples have been involved in parenting, it has been through joint custody or visitation rights. Other case law has maintained that sexual orientation should not interfere with visitation rights. Other courts have opposed custody for gay and lesbian parents on the ground of possible problems posed for the children at school and in the community by having two same-sex parents.

The three principal routes to lesbian or gay parenthood are where (a) one member of the couple brings one or more children from a heterosexual marriage, (b) the couple adopts, and (c) one partner (among lesbian couples) becomes pregnant through donor insemination. Evidence reported from a number of studies has not shown that children living with gay or lesbian couples face the problems in adjustment that has been presumed by some courts. Whether considered from the perspective of teachers, parents, or children, findings regarding the adjustment of children of gay or lesbian parents point to relatively few problems. Patterson's report suggested that children of lesbian mothers reported somewhat greater stress than did children of heterosexual mothers, but they also reported an enhanced sense of well-being and comfort with their own lives. Patterson's review of studies in this area showed that children of gay or lesbian parents have the expected social ties during middle childhood and report few problems at school as a result of having two same-sex parents.

Children of gay fathers have somewhat greater difficulty in accepting their parents' sexual orientation than do children of lesbian mothers, although the nature of children's acceptance depends on their age at the time of parental disclosure. School-aged children have a somewhat easier time than do adolescents in making sense of parental sexual orientation for their own lives. Patterson observed that early adolescence may be a particularly difficult time for children to learn about their parents' alternative sexual orientation. Where problems with having gay parents have been reported, most notable among gay couples with children either adopted or brought from prior marriages, the extent of problems reported by the children is related to the degree of comfort that their same-sex parents have with their own role as gay men and fathers.

Child care and division of labor in doing routine household tasks appear to be more of a problem among gay couples than among lesbian couples. Too often in our society, child care has been considered the work of women and fathers have served as children's primary caregivers less often than have mothers. Gay fathers often try to compensate for any hardship their sexual orientation may pose for their children. They characterize themselves as both warm and nurturing as well as more likely than heterosexual fathers to set limits. Among both gay and lesbian couples, the two partners are more likely to be supportive and involved in child care than is the case among heterosexual couples, where child care is most often seen as primarily the responsibility of the wife and mother. Gay fathers, in particular, have had to work so hard to attain custody or to adopt that they are particularly likely to be involved in parenthood.

Gay fathers and lesbian mothers sometimes report anecdotally that they are concerned that their children should adopt gender-appropriate roles and engage in gender-stereotyped activities that conflate gender and sex roles but provide assurance that their offspring will not adopt a gay or lesbian identity as a reflection of their own still unresolved feedings about their gay or lesbian identity. In the past, gay fathers expressed greater concern than did lesbian mothers regarding the impact of their sexual orientation on their children's development. Although the offspring of gay and lesbian parents are not more likely than the

offspring of heterosexual parents to report a gay or lesbian sexual orientation, their parents are particularly concerned that their offspring adopt gender-stereotyped interests and not show evidence of a same-sex sexual orientation. This somewhat exaggerated concern reflects their awareness of social bias, their own experiences of growing up in the midst of a hostile climate at school and in the community, and their own internalized anti-gay prejudice.

7. CONCLUSION

Understood as same-gender sexual orientation and the consequent assumption of a sexual identity as gay or lesbian, homosexuality has assumed a changed significance in postwar society. Formerly a furtive identity that necessarily required "covering" in a conservative heterosexual society, urban-living lesbians and gay men (and, to some extent, even those living in smaller towns) have enjoyed increased social acceptance and diminished prejudice over the past two decades or so. In turn, this enhanced acceptance has led to greater ease among youth disclosing to friends and family their gay or lesbian sexual identity and has permitted gay and lesbian adults to enjoy their lives with friends, family, and children, comparable to the lives of their heterosexual counterparts. Much of what is known about the impact of alterative sexual identity on personal development across the course of life has been based on these generations growing up in the midst of prejudice. Challenges for the next period of study include long-term study of the adjustment of children of gay and lesbian parents as they reach adulthood and the impact of sexual orientation on the adjustment to aging as "baby-boomer" lesbians and gay men reach later life.

See Also the Following Articles

Gender and Culture ■ Gender Role Development ■ Homophobia ■ Prejudice and Discrimination ■ Risk-Taking and Sexual Behavior ■ Sexual Behavior ■ Sexual Behavior and Culture ■ Stereotypes

Further Reading

Bailey, J. M., & Pillard, R. (1991). A genetic study of male sexual orientation. *Archives of General Psychiatry*, *48*, 1089–1096.

Cabaj, R. P., & Stein, T. S. (1996). *Textbook of homosexuality and mental health*. Washington, DC: American Psychiatric Press.

Chauncey, G. (1994). *Gay New York: Gender, urban culture, and the making of the gay male world, 1890–1940*. New York: Basic Books.

Chodorow, N. (1992). Heterosexuality as a compromise formation: Reflections on the psychoanalytic theory of sexual development. *Psychoanalysis and Contemporary Thought*, *15*, 267–304.

Cohler, B. (in press). *Writing desire: Generation and life-story among men who seek sex with other men*. Madison: University of Wisconsin Press.

Cohler, B. J., & Galatzer-Levy, R. L. (2000). *The course of gay and lesbian lives: Social and psychoanalytic perspectives*. Chicago: University of Chicago Press.

Drescher, J. (1998). *Psychoanalytic therapy and the gay man*. Hillsdale, NJ: Analytic Press.

Friedman, R. C., & Downey, J. I. (2002). *Sexual orientation and psychoanalysis*. New York: Columbia University Press.

Group for the Advancement of Psychiatry. (2000). *Homosexuality and the mental health professions: The impact of bias*. Hillsdale, NJ: Analytic Press.

Hamer, D. H., Hu, S., Magnusson, V., Hu, N., & Pattatucci, A. (1993). A linkage between DNA markers on the X chromosome and male sexual orientation. *Science*, *261*, 321–327.

Herdt, G. H. (1997). *Same sex, different cultures*. Boulder, CO: Westview.

Herek, G. M. (1993). The context of anti-gay violence: Notes on cultural and psychological heterosexism. In L. D. Garnets, & D. C. Kimmel (Eds.), *Psychological perspectives on lesbian, gay, and bisexual experiences* (pp. 89–108). New York: Columbia University Press.

Malyon, A. (1982). Psychotherapeutic implications of internalized homophobia in gay men. *Journal of Homosexuality*, *72*, 59–69.

Murray, S. O. (2000). *Homosexualities*. Chicago: University of Chicago Press.

Patterson, C. J. (2000). Family relationships of lesbians and gay men. *Journal of Marriage and the Family*, *62*, 1052–1069.

Plummer, K. (1995). *Telling sexual stories: Power, change, and social worlds*. New York: Routledge.

Posner, R. (1992). *Sex and reason*. Cambridge, MA: Harvard University Press.

Purcell, D. W. (1998). Current trends in same-sex marriage. In R. P. Cabaj, & D. W. Purcell (Eds.), *On the road to same-sex marriage* (pp. 29–40). San Francisco: Jossey–Bass.

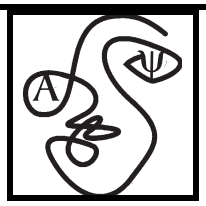
Savin-Williams, R. C. (2001). *Mom, dad, I'm gay: How families negotiate coming out*. Washington, DC: American Psychological Association.

Stein, E. (1999). *The mismeasure of desire: The science, theory, and ethics of sexual orientation*. New York: Oxford University Press.

Sullivan, A. (Ed.). (1997). *Same-sex marriage: Pro and con*. New York: Vintage Books.

Sullivan, A. (1998). *Love undetectable: Notes on friendship, sex, and survival*. New York: Knopf.

Warner, M. C. (1999). *The trouble with normal: Sex, politics, and the ethics of queer life*. New York: Free Press.



Human Rights

Willem Doise
University of Geneva, Geneva, Switzerland

1. Institutional Definition of Human Rights
2. The Ambiguous Message of Social Psychology
3. Human Rights as Normative Social Representations
4. Research Findings
5. Putting Things in Perspective
Further Reading

GLOSSARY

collective rights Particular rights based on membership in cultural, religious, linguistic, or national groups. The main function of such rights is to maintain and further links between members of a community as well as to preserve their lifestyle.

contextualization of rights Processes that actualize a more or less broadening or restricting definition of rights that are deemed to be relevant in a given situation.

human rights Universal rights that all individuals should enjoy regardless of their belonging to different social groups or categories. Individuals are invested with these rights by the mere fact of being a member of the human species. The Universal Declaration of Human Rights, voted by the General Assembly of the United Nations in 1948, is the official document stating these rights.

normative social representations Representations that specifically allow humans to evaluate their relations and interactions. All social representations imply a normative aspect because they are beliefs embedded in systems of norms.

social representations Organizing principles of symbolic relationships between individuals and groups. They imply common understanding of social reality, organization of individual positioning toward issues at stake in society, and anchoring of individual variations of positioning in social realities.

Human rights are defined as normative social representations embedded in institutional juridical definitions. Conclusions from psychological studies on matters of rights enforcement are ambiguous. Research findings show that human rights can be studied as normative social representations implying a degree of common understanding across cultures together with organized differences within and between cultures. Important factors in modulating individual positioning in the realm of human rights are experiences of social conflict and injustice, beliefs about the efficiency of various social actors to have rights enforced, and attitudes of liberalism or collectivism.

1. INSTITUTIONAL DEFINITION OF HUMAN RIGHTS

The Universal Declaration of Human Rights, adopted on December 10, 1948, by the General Assembly of the United Nations Organization, serves as the main institutional reference for defining human rights (HR). The declaration comprises six groups of articles. The first group of articles (Nos. 1 and 2) enunciates the basic principles of equality, freedom, and dignity; the second group (Nos. 3–11) focuses on rights of the individual,

©2004 Elsevier Inc.
All rights reserved.

such as security of the person and equal protection by the law; the third (Nos. 12–17) concerns rights relative to interindividual relations, among which there are freedom of movement and the right to found a family; the fourth (Nos. 18–21) involves civil rights, such as freedom of expression or equal access to public service; the fifth group (Nos. 22–27) deals with economic and social rights, such as social security and the right to rest and leisure; and the sixth (Nos. 28–30) covers societal rights relative to international law and order as well as to duties to the community.

When the declaration was drafted, it was not at all obvious that eventually all groups of articles would be retained by the vote of the general assembly. Fierce controversies opposed members favoring exclusively a declaration of individual judicial rights together with public and civil rights and members more concerned about providing individuals with socioeconomic rights, such as rights to a decent standard of living, education, and health care. This antagonism is still acute and is reflected in the contents of the two additional covenants of 1966 that have defined a more legal status for rights defined in the declaration, one covenant bearing on civil and political rights and the other on economic, social, and cultural rights. These resulted in procedures aimed at guaranteeing respect of specific rights, at least among the countries that signed them. Not all the countries that signed the universal declaration signed both additional pacts. Hence, the universal declaration remains the most universal and comprehensive reference in the realm of human rights.

Half a century after its adoption, it is more than evident that enforcement of the rights stipulated in the declaration is far from being guaranteed in many countries throughout the world. Reasons for this lack of HR respect are numerous and of diverse nature: political, economical, cultural, institutional, and also psychological.

2. THE AMBIGUOUS MESSAGE OF SOCIAL PSYCHOLOGY

Psychological dispositions without doubt influence the way individuals take into consideration the rights of others. In order to study such influences, one could borrow a lot of useful suggestions from studies on prosocial behavior. However, in the present context, human rights enforcement is considered, above all, a matter of institutions warranting rights of individuals. The functioning of institutions involves individuals in that they participate in specific relationships. Therefore, I privilege here social psychological analyses of such functioning rather than analyses in terms of individual differences.

Lessons of social psychology in relation to human rights are often lessons in skepticism. In research performed by Stanley Milgram in 1974, participants went ahead in delivering what they thought to be dangerous electric shocks to a person refusing to be further shocked or who was already unconscious. In a series of critical replications published by Meeus and Raaijmakers in 1987, it was shown again that for the sake of an experiment on stress, many people prevented an unemployed man from being hired. Of course, unknown to the experimental subjects in these experiments, the unemployed man was an actor.

Another lesson of skepticism is that in order to uphold a basic belief in a just world, many people often denigrate and chastise other people for the simple reason that they suffer. According to Melvin Lerner in his monograph published in 1980, a just world belief, seemingly very widespread in Western societies, would imply that remuneration and punishment, positive sanctions, and negative sanctions are not distributed randomly in this world. If one suffers, there should be a reason, and therefore people who suffer are often denigrated and considered responsible of their own fate. Literature on victimization, as summarized by Janoff-Bulman and Hanson Frieze in 1983, is based on a similar hypothesis. This trend of research can be illustrated by a concrete example: Witnesses of a rape scene, and even the victim of such an aggression, often tend to blame the victim for having an important part of responsibility in the aggression.

Lerner puts forward the idea that the fundamental need for justice could easily turn to discriminatory practices. In order to satisfy one's own needs as well as those of one's next of kin in an appropriate manner, it would be necessary to keep one's distance from a large part of humanity that cannot satisfy its hunger and that lives in inequality and injustice. To interact with such people outside strictly regulated or even discriminatory terms is looked upon as risky behavior endangering the progress of justice at home. It would be precisely those people who subscribe most to the concept of a just world who would best accept institutionalized patterns of discrimination.

Note, however, that the principle of a just world can also be put forward to explain altruistic behavior. It would thus account for feelings of indignation felt when witnessing some blatant injustices and could

even occasionally be at the base of "heroic" attempts made to redress these torts.

Research on submission to authority as well as that on the perverse effects of a belief in a just world are examples of the various trends of research that justify an attitude of skepticism about the often optimistic vision maintained by human rights supporters. Another important point to be added to this rather grim picture is the diffusion of responsibility described by Latane and Darley in 1970. Their studies aim to explain why individuals often do not intervene when confronted with an emergency, such as when somebody is attacked by another individual or falls in a state of unconsciousness on the street. A robust finding is that the presence of other bystanders reduces to an important extent the willingness of individuals to intervene in such a situation. Apparently, people feel less concerned to intervene in a situation of emergency when others can also intervene. Theoretically, one can suppose that this is often the case in situations of human rights violations and that therefore a reaction of "it is not my responsibility to intervene" is very likely to be manifested. Taking the rationale further, individuals confronted with massive and institutionalized HR violations may think that their own role in redressing these violations is utterly insignificant. Fortunately, when broadening the realm of research covered by psychologists in order to include research on the social development of children and adolescents, indications favoring more optimistic conclusions start to emerge.

What characterizes modern thinking is the refusal that traditional bodies such as churches, authorities, or charismatic personalities decide what is right or wrong, what is good or bad. It is this refusal that necessitates the building of a constructivist concept of knowledge and ethics. This conception of modernity oriented the groundbreaking research by Jean Piaget published in 1932 on *The Moral Judgment of the Child*. His aim was to study the transition of morals among children based essentially on the approval of rules decreed by others to more autonomous morals allowing for contractual changes of these rules through discussion and agreement between parties involved. According to Piaget, participation of children in interaction with equals makes such a transition from heteronomous to autonomous morality possible.

To consider that this developmental hypothesis by Piaget has been empirically authenticated would be exaggerated. Nevertheless, it embodies a general concept that has been at the root of other investigations too numerous to list. Only a brief presentation of the stages of moral development described by Lawrence Kohlberg in 1984 is provided. Indeed, Kohlberg has refined, to a large extent, Piaget's distinction between the two kinds of morality. He describes no less than six stages in the moral judgment of individuals. It is important to briefly characterize these stages, especially because the final two offer foundations for a constructivist concept of human rights ideas.

The first four stages are not directly related to human rights. The first one is essentially governed by constraints issuing from rules and sanctions: The avoidance of punishment and respect of authority are considered as ends in themselves. During the next stage, an individualistic orientation still predominates, but it already includes an instrumental concept of trading and reciprocity. The respect of the interest of others is used to better guarantee one's own interests. The third stage represents conformism and seeking of social recognition. Attitudes such as to be well intentioned, to respect others and to keep one's promises, and to remain faithful in allegiances give rise to morals illustrated by the statement "Do not do unto others what you would not have others do unto you." The fourth stage is based on the principle of duty—the need to follow laws, rules, and conventions in order to keep the social system functioning well. A majority of adults often seem to reason following principles of the third and fourth stages.

The fifth stage sees the explicit intervention of a relativist attitude by acknowledging that many individual values and opinions are characteristic of given groups or societies and that other social contracts could set a higher value on different attitudes. The problem of the universality of specific values such as freedom or respect for life is explicitly put forward and a great significance is assigned to the consensual definition of a procedure allowing the fixing of rights and values to be respected by all.

The sixth and last stage is hardly encountered elsewhere than in writings of specialists such as philosophers and human rights militants. Universal ethical principles should supersede rules and conventions when they are in conflict. Following Immanuel Kant, people are considered an end in themselves. The moral point of view is the one that all human beings should adopt toward each other as free and autonomous subjects. Reasoning characteristic of this stage should take into consideration the point of view of each person susceptible of being affected by consequences of a decision to be made; such reasoning is based on the hypothesis of reversibility, according to which everyone should be able to put themselves in the others' positions.

This is not the occasion to start a technical debate on the theoretical value of Kohlberg's stage model in the frame of current developmental psychology. It would be equally wrong to pretend that a person would necessarily fit into only one stage, as well as to pretend that their order of succession in individual development would occur at random. For present purposes, the important point is that an idealized conception of human relationships has also been the object of psychological studies.

In *Minding the Law*, A. G. Amsterdam, an expert in U.S. constitutional law, and Jerome Bruner, whose contributions to the psychology of child development, cognition, education, language, and culture are well known, analyze a series of decisions by the U.S. Supreme Court related to racial discrimination and the equal protection clause of the Fourteenth Amendment. Studying relevant documents and opinions of the Supreme Court in terms of category processes, narratives, rhetorics (conceived as social influence devices), and cultural dynamics, the authors show how important variations occur over time. The cases they comment on more extensively deal with the rendition of a runaway slave (*Prigg v. Pennsylvania*, 1842), racial segregation in railroad transports (*Plessy v. Ferguson*, 1896), racial desegregation in public schools (*Brown v. Board of Education*, 1954), the death sentence of an African American (*McCleskey v. Kemp*, 1987), and restrictions on desegregation in public schools (*Freeman v. Pitts*, 1992, and *Missouri v. Jenkins*, 1995).

The main conclusion of the study on historical variations over more than a century and half in the decisions by the Supreme Court is framed by Amsterdam and Bruner in the following terms: "This trail can be seen as marking the fall, the rise, and the fall again of a brightly egalitarian and progressive American Creed—or as marking the rise, the fall, and the rise again of an equally potent, darkly suspicious American Caution" (p. 16). Indeed, as others before them, the authors conclude that a basic component of American culture is the principles proclaimed in the Declaration of Independence, but they also conclude that the evolution of this culture cannot be understood if one does not admit the intervention of an equally potent principle restraining and limiting these rights for long periods. In discussing the links between these two fundamental aspects of American culture, they echo ideas of Lerner: "So the preservation of the American Creed requires that We (sic) create a category of people who are essentially and immutably different and inadmissible" (p. 264).

Hence, taking into consideration various research traditions, there seems to be a double message: Psychologists deliver lessons of optimism as well as lessons of pessimism concerning matters related to human rights. According to Piaget and Kohlberg, the ideal of human rights would be, in a certain way, a spontaneous product of democratic interaction patterns. According to Milgram, citizens of democratic states can easily be brought to violate fundamental rights of others, and according to Lerner the proper belief in a just world may be used to justify practices of social discrimination as well as attempts to restore justice. In their contribution to the study of U.S. law and culture, Amsterdam and Bruner conclude that adhesion to an optimistic egalitarian credo in a liberal democracy can be accompanied by an equally strong belief that rights of certain groups of people should be severely limited.

At first sight, these messages carry insurmountable contradictions that should resist integration within a same theoretical framework. However, studying human rights as normative social representations can allow for a better understanding of these seemingly conflicting lessons and offer a possibility of integrating them in a common explanatory frame.

3. HUMAN RIGHTS AS NORMATIVE SOCIAL REPRESENTATIONS

Multiple relations interconnect members of the human species directly or indirectly. Between the different groups constituting mankind circulate not only genes, viruses, pollutants, epidemics, consumer goods, or money but also ideas and institutions. What inhabitants of one country practice affects to varying degrees the fate of people abroad, their way of farming land, their state of health, their access to all kinds of resources, and the functioning of their institutions. It is difficult to imagine that a group of humans would not be affected to some extent by the way of life of other groups. In this sense, apartheid does not exist, but on the other hand, there is no perception, no definition, nor exhaustive analysis of the numerous ties that bind humans together. Even if globalization is a new term for naming that general interdependence, it does not mean that the phenomenon can be described precisely and fully understood. It is also true that general terms, such as globalization, can be used to conceal the obvious reality of exploitation or of negative interdependence. Such realities, even though they are part of a larger relational field, exercise their own noxious effects in many ways.

Interaction and communication between humans initiate symbolic representations, social norms, and contractual principles that often remain implicit. When we enter into a relationship with other persons, we know that the fate of participants in such a relation will be affected by their interaction, in some measure, within certain limits, at a certain cost. Normative representations exist on what these effects should be. Multiple forms of interdependence exist, characterized by all sorts of differences in status, purpose, interdependency, and formality. Various models of acceptable relationships, prototypes of fair and just situations, and contractual principles govern evaluations of these relationships. Such evaluative criteria have been built in multiple interactions, they are culturally defined, and their application can be guaranteed by institutions. Some of them are culturally specific, whereas others are spread over different cultures.

Human rights are such evaluative principles or normative social representations that could allow, at least at the level of the intention, humans to evaluate and to organize their relations and interactions. For historical (i.e., economical, political, military, religious, and scientific) reasons, Western societies organized relationships within national and cultural boundaries as well as across them. Many encounters with foreign cultures were sources of continuous conflict, and much irremediable harm has been inflicted in the frame of these contacts. Nevertheless, revolutionary ideas about universal human rights that were initially formulated as reactions against the arbitrariness of established powers (at the occasion of the American Declaration of Independence in 1776 and the French Declaration of the Rights of Man and Citizen in 1789) were thereafter progressively invested with new meanings. These meanings were elaborated by diverse and sometimes antagonistic militant movements in the realms of antislavery, antiserfdom, antiracism, and anticolonialism causes; by movements defending rights of workers, women, cultural minorities, aboriginal peoples, wounded soldiers, displaced persons; and by movements furthering right claims of many other social categories. The visions of human rights elaborated by these movements, and often forged at the expense of much effort and struggle, became institutionalized in the Universal Declaration of 1948, which is to be considered according to its preamble

> As a common standard of achievement for all peoples and all nations, to the end that every individual and every organ of society, keeping this Declaration constantly in mind, shall strive by teaching and education to promote respect for these rights and freedoms and by progressive measures, national and international to secure their universal and effective recognition and observance, both among the peoples of Member States themselves and among the peoples of the territories under their jurisdiction.

Human rights are clearly normative social representations, even if they are also important political and judicial realties.

It is generally accepted that in interactions with others, independently from their origin, the principles adopted in the Universal Declaration compel the respect of corporal integrity of these others, their liberty and their dignity, their access to vital resources, and their integration in a societal order guaranteeing protection against arbitrary decisions. Logically, in order to respect commitments taken by public authorities, citizens of many countries should only involve themselves in relationships respecting these principles. However, it is a fact that in numerous situations basic rights of individuals are not respected, as evidenced by economic conditions imposed on Third World countries or when some people, even in First World countries, are deprived from basic health care or from subsistence-level resources.

Here, a huge field of research opens. If human rights are indeed normative social representations, social psychologists—at least those involved in social representation studies—have theoretical and methodological tools at their disposal to study them.

4. RESEARCH FINDINGS

Social representations (SRs) are defined as organizing principles of symbolic relationships between individuals and groups. A first assumption on which this definition is grounded is that various members of a population under study share common views about a given social issue. SRs are generated in systems of communication that necessitate common frames of reference for individuals and groups participating in these communications. Hence, studying human rights in SRs terms always implies a search for common reference systems and for their organization. To what extent does the Universal Declaration offer references that are common to different populations under study?

Mainly, three studies are discussed here because they deal with SRs of HR in a transcultural setting. Two of them are extensively described in the 2002 monograph by Doise.

The first study derived from an interdisciplinary interview study on representations of violations. In a

questionnaire study, students aged 13–20 years in four countries (Costa Rica, France, Italy, and Switzerland) were invited to answer 21 items presenting various situations involving violations or limitations of individual rights. Some of these situations (e.g., racial discrimination, imprisonment without trial or legal assistance, and starvation) can easily be referred to classical definitions contained in the Universal Declaration. Other situations, dealing with the rights of children or with family affairs, are less explicitly related to the declaration. Some situations dealing with economic inequality or health matters (e.g., prohibition of smoking or hospitalization in case of contagious illness) are not covered by official definitions of rights.

The results were clear: Across countries generally the same situations are more often considered to be human rights violations. These findings have since been confirmed by research in the United States and Romania.

In a second transcultural study, the complete text of the declaration was presented to students of 35 countries. Respondents (university students in psychology, law, science, social work, and various other fields from the five continents) were asked to answer questions about personal involvement, agreement and efficacy, as well as governmental efficacy for each of the 30 articles of the declaration. A hierarchical cluster analysis designed to accentuate intraclass homogeneity was performed on indices of the Euclidean distances between articles based on the sums of responses to the eight scales for each article. This analysis resulted in the division of the articles into two main classes, which in turn were divided into subclasses. These subclasses showed an almost complete correspondence with the groups of articles described at the beginning of this article.

A subclass of the first main class aggregates the more social articles, relational rights (group 3, except article 14 about asylum), public liberties (group 4), and articles dealing with economic, social, and cultural rights (group 5). The other subclass aggregates the three articles referring to the more basic individual rights (protection from torture and slavery and right to life; articles 3–5 from group 2).

The second main class can be subdivided into three classes: a subclass of judicial individual rights (remaining articles of group 2) together with the right to asylum (article 14), basic principles (group 1), and the three articles concerning societal order (group 6).

In each national group, respondents show greater adherence to the basic and social rights than to the rest of the rights. These results clearly support the idea of a common organization of responses in various countries. Such a common organization was also observed by Spini in a follow-up study using other samples of respondents from 16 countries belonging to five different sociogeographical areas.

A third study was carried out by Herrera and Doise in the province of Quebec, Canada. Participants in this study were French-speaking and English-speaking Quebecois and members of the aboriginal Innu nation (Montagnais). Unlike the previous studies, the questionnaire used in this study started with two different open-ended questions asking respondents to mention the human rights they know and also the specific rights claims of respectively the Quebecois, the Quebecers, and the autochthones according to the ethnic group membership of the three samples under study. For instance, for the Innu sample, this question was worded as follows: "According to you, which are the five most important rights of autochthonous people that would do them justice?"

Analysis of answers for the three groups shows that most of the rights that were mentioned as examples of HR could indeed be related to the Universal Declaration of Human Rights. There was also a high convergence between evocations of HR by the three groups: The four groups of rights most often referred to by the three samples were public freedoms and political rights, basic human rights principles, fundamental individual rights, and economic social and cultural rights.

Such a convergence was not observed for the evocations of collective rights (i.e., rights of the Quebecois or of the Aboriginal people also investigated in this study). Even if almost all rights mentioned as collective rights could be related to both the Universal Declaration of Human Rights and the Draft Declaration of Rights of Indigenous People, the rights more often mentioned by respondents of the three groups were not the same. The members of the Innu sample mentioned rights to dispose of their own territory and resources, fundamental collective rights such as the right to autonomy, rights to the preservation of their traditions and language, and rights to self-government and specific indigenous laws more often than other categories of rights. Of the rights contained in the Draft Declaration of Rights of Indigenous Peoples, only the rights concerning language and tradition figured among the rights most often quoted by both other samples.

Findings of this study again showed to what extent ideas about human rights are shared across group boundaries, although collective rights are interpreted in a more diverse way.

Social representation theory does not imply that individuals sharing common references hold the same positions. It is therefore important to investigate differences in individual positioning. Modulators of positioning are beliefs about personal efficacy and the efficacy of institutions (e.g., governments) with respect to human rights. Human rights positions also bear a relationship with value choices. Values as measured by Rokeach's terminal values are considered general to the extent that they supposedly organize symbolic relationships within a social environment. Given the historical origins of human rights ideas in conditions of discrimination and conflict, it is also expected that HR representations will be modulated by individual experiences of injustice and of conflict between social groups and categories.

A typology was constructed of respondents in the study with the 30 articles of the declaration. Four groups were retained: A group of advocates and a group of skeptics had relatively the highest and lowest scores on scales concerning general importance of HR and personal and governmental involvement in enforcing rights. A group of personalists considered that HR concerned them personally but that it was not easy for governments to do something for their enforcement, whereas governmentalists considered themselves rather powerless for having HR respected but considered the government as relatively more effective. Strong support for the values of universalism and social harmony was found to be systematically related to HR advocacy. Intense experience of collective injustice, together with less concern for personal happiness, was characteristic of personalists rather than of governmentalists. It was also found that skeptics were relatively more numerous in Japan and India, and personalists were more often found in countries with serious human rights problems according to the 1992 assessment published by C. Humana and with human development problems according to the 1996 ratings by the United Nations Development Programme. Governmentalists were relatively more numerous in developed countries or countries that changed recently to a democratic regime.

Individual positioning in the study on violations of HR was also investigated. For the respondents in the initial four-country study and the Romanian sample, a cluster analysis extracted three groups of individuals according to the similarity of their responses. The ranking of violations for the two main clusters clearly overlapped the ranking found for the total population. However, these two clusters could be differentiated mainly on situations related to family affairs. Obliging children to forego schooling or to attend mass, parental child beating, and preventing one's spouse (wife or husband) from going out alone were less clearly perceived as human rights violations by a group of respondents sharing a restricted definition of human rights, as were the obligation for women to veil their faces and the refusal to rent an apartment to a black person or a gypsy family. Respondents adhering to a more extended definition considered all these cases to be violations of HR. On the other hand, responses given by members of a third cluster were clearly atypical, probably reflecting a relative disinterest for the HR cause, at least in the manner they were institutionally defined.

How were individual variations in social positioning toward human rights violations linked to positioning in other societal realms? In order to investigate these relations, five additive scores were computed for each individual on a series of anchoring variables, namely the number of governmental actions considered unacceptable (e.g., death penalty and suppression of elections or of television broadcasting critical of government), the number of queries made by government and business firms considered to be unfair (e.g., concerning health and militancy in unions), the number of official regulations (e.g., regarding vaccination and compulsory military service) interpreted as infringement of individual freedom, the number of individual actions reported as unacceptable (e.g., bribery, hiding an escaped convict, and giving in to blackmail), and the number of agreements with fatalistic assertions (e.g., there will always be rich and poor, good and bad people, and nations that do not respect HR).

These anchoring variables were then used as independent variables in a discriminant function analysis with the three clusters of the typology as dependent variables. The two functions were statistically significant. The first function depicts the specific anchoring of the cluster of adherents to a restricted definition. In comparison to members of the other two clusters, individuals adhering to a conception that adheres more closely to the institutional definition of human rights cultivate a relatively more fatalistic conception of social reality while expressing broader tolerance toward governmental interference and various social regulations. One can clearly identify the origin of a restricted definition of rights: increasing power given to the government and acceptance of various social regulations in the context of a fatalistic conception of social relations. The second function merges an extended definition with refusal of various governmental and managerial infringements, with a higher tolerance for questionable individual deeds and a better

adhesion to various social regulations. In other words, this conception focused on protecting the individual—even while occupying a marginal social position—against authorities while accepting their regulatory role. The location of the atypical or unconventional position on each of these two functions reveals the complete opposition of its anchoring to those described previously.

In the Quebec study by Herrera and Doise, some HR were mentioned as collective rights and vice versa. A factor analysis shows that answers to both questions are interrelated in various ways, but that individual scores on the factors are again related to the way respondents perceive the actual and possible contribution of different social actors in having human and collective rights respected as well as to their experience of discrimination and rights violations.

A nonhierarchical cluster analysis classifies respondents into three groups. The position of members of a first group can be called one of liberalism. Tenants of this position are observed to adhere above all to the basic principles of human rights and to public freedom rights. The position shared by members of a second group corresponds to the cause of sovereignty that privileges mainly cultural and linguistic rights, but this position also implies an awareness about self-government issues and problems related to socioeconomic rights compatible with HR principles. These two profiles were more characteristic of nonaboriginal French-speaking as well as English-speaking participants who seemed equally aware that in their sociocultural context important issues in defining HR and collective rights are those of liberalism and sovereignty. Self-determination, especially in matters concerning the use of land and resources and self-government, characterizes the position of a third group of respondents. As could be expected, the latter profile was characteristic of the aboriginal sample.

The differences of the SR approach and other social psychological approaches in HR studies are difficult to assess because few psychologists have researched HR. In a 1995 study, Diaz-Veizades *et al*. shared some of the concerns mentioned in this article because they also used the Universal Declaration of Human Rights in order to study the organization of individual positioning as well as anchoring of this positioning. However, they undertook important adaptations of the original document, reformulating the 30 articles of the declaration into 116 more concrete items, such as "If a person does not make enough money to support his or her family adequately, the family should be aided by the government," "There are times when people should be kept from expressing their opinion," "Men and women should have equal rights in a marriage," and "A person's home is his or her 'castle' and should not be interfered with by others."

For each item, their respondents (mainly North American college students) expressed their degree of agreement on a 7-point scale. However, the authors report that many items "had very low variance because of high rates of endorsement, so they could not correlate highly with other items. After these low-loading items were deleted, 38 items were reanalyzed using an iterated principle factor analysis." Clearly, we are confronted here with a logic different from the one adopted in this article. Diaz-Veizades and colleagues favored the study of systematic interindividual variations to the detriment of analyzing the common meaning aspect of SRs. When eliminating the most consensual items of their questionnaire, they could retain four factors for the remaining items that represented only one-third of all items presented to respondents. A first factor, called the social security factor, concerned access or entitlement to an adequate standard of living (e.g., food, housing, and medical care). A second factor, civilian constraint, dealt with the acceptability of limiting individual civil and political rights. The theme tying items together for the third factor was that of equality, evidenced most clearly by items dealing with equal access to basic rights for all individuals regardless of race, gender, or beliefs. Finally, items with highest loading on the fourth factor involved privacy issues. Diaz-Veizades and colleagues also analyzed anchoring patterns in showing, for instance, that respondents adhering more to a civilian constraint conception of HR were also those who obtained higher scores on a nationalism scale and lower ones on internationalism and civil liberties scales; their political preferences were more likely to favor Republicans over Democrats as opposed to the preferences of adherents to a social security conception, who favored Democrats and had higher scores on internationalism.

5. PUTTING THINGS IN PERSPECTIVE

An important issue that preoccupies colleagues from anthropology and cultural psychology is the problem of the universality of HR, and some of them envisage the possibility that HR are merely an expression of

Western ethnocentrism. Indeed, ethnocentric use of HR was experimentally evidenced by Moghaddam and Vuksanovic in 1990. In a first study, they asked undergraduate students in Montreal to answer a questionnaire about HR issues. The questionnaire is identical in three conditions, except for the fact that in a first condition the context referred to was Canada, in a second condition it was the Soviet Union, and in a third condition it was the Third World. Typical items were as follows: "All forms of censorship should be done away with in [Canada/Soviet Union/Third World] societies" and "Everyone in [Canada/Soviet Union/Third World] societies should have access to free health care." Overall, support for HR was stronger in the Soviet Union and Third World contexts than in the Canadian context. In a second study, three scenarios were used concerning a television news anchor woman who had been fired from her job, a member of a terrorist group who had been tortured, and a female shoplifter, again presented in three different contexts. As in the first study, support for HR was stronger in the conditions Soviet Union and Third World than in the condition Canada. Such results tend to confirm the opinion of critics who consider HR as a Western export article, considered useful for others but less so for Westerners themselves.

The previous discussion suggests that the HR issue is indeed a highly normative one. Individual stances are embedded in value choices and orient judgments about social reality also in the realm of international relations. In one's own environment, strong convictions about values are not necessarily applied in a consistent way. Clémence and Staerklé explored this issue and their findings are reported in the monograph by Doise. When comparing judgments on rights limitations reported in the realm of privacy, free education, or the death penalty, they found that Genevan participants in their study very strongly adhered to the relevant HR principles, giving average scores higher than 6 on a 7-point scale. However, scores dropped significantly when it was a matter of indicating their degree of acceptance (1, acceptable; 7, unacceptable) of violations of these rights in cases of a search without a warrant of a home by a police officer, refusal of free primary schooling to handicapped children, or the death penalty for minors. Furthermore, when these limitations of individual rights were merely reported as newspaper items ("faits divers"), their degree of rejection was on average 4.94, whereas when they were explicitly reported as extracts of HR reports, the average degree of unacceptability decreased by almost 1 point on the scale to 4.01. It seems as if references to HR for events in Western countries were less pertinent and prevented people from going into a more detailed analysis of the situation. Paradoxically, the absence of an explicit invocation of HR implies that individuals pay greater attention to reprehensible aspects of acts occurring in their own environment and, as a consequence, find them less acceptable. On the contrary, when general HR principles are explicitly evoked, concrete characteristics of events happening in one's own country are not taken into account, and in this way people can moderate their judgments.

The fact that HR are to be considered normative social representations invests them with very strong value connotations. As a consequence, in some situations violations of rights are considered less important when they happen in one's country, whereas HR violations far abroad are strongly condemned.

Without doubt, variations in enforcement of HR exist across countries. However, differences in beliefs related to the actual enforcement of HR in one's own country also exist between inhabitants of Western democratic countries. The principal component of a normative social representation will certainly remain its insufficient realization in many social contexts. Otherwise, it would no longer have reason to exist. It is precisely the awareness that rights are violated that often makes people claim that these rights should be enforced—that their enforcement has to become part of social reality.

Further Reading

Aboriginal and Torres Strait Islander Commission (1996). *United Nations Draft Declaration on the Rights of Indigenous Peoples*. Canberra, Australia.

Amsterdam, A. G., & Bruner, J. (2000). *Minding the law*. London: Harvard University Press.

Bierhoff, H. W. (2002). *Prosocial behaviour*. New York: Taylor & Francis.

Black, C. L., Jr. (1997). *A new birth of freedom. Human rights named and unnamed*. New York: Grosset/Putnam.

Diaz-Veizades, J., Widaman, K. F., Little, T. D., & Gibbs, K. W. (1995). The measurement and structure of human rights attitudes. *Journal of Social Psychology*, *135*, 313–328.

Doise, W. (2002). *Human rights as social representations*. London: Routledge.

Doise, W., Clémence, A., & Lorenzi-Cioldi, F. (1993). *The quantitative analysis of social representations*. Hemel, Hempstead: Harvester Wheatsheaf.

Finkel, N., & Moghaddam, F. (2004). *The psychology of rights and duties: Empirical contributions and normative commentaries*. Washington, DC: APA Books.

Herrera, M., & Doise, W. (2004). Social representations of human and collective rights in three sociocultural groups of Quebec. *New Review of Social Psychology, 3*.

Humana, C. (1992). *World human rights guide* (3rd ed.). New York: Oxford University Press.

Janoff-Bulman, R., & Hanson Frieze, I. (1983). A theoretical perspective for understanding reactions to victimization. *Journal of Social Issues*, *39*, 1–17.

Kohlberg, L. (1984). *The psychology of moral development*. San Francisco: Harper & Row.

Lauren, P. G. (1998). *The evolution of international human rights. Visions seen*. Philadelphia: University of Pennsylvania Press.

Lerner, M. J. (1980). *The belief in a just world. A fundamental delusion*. New York: Plenum.

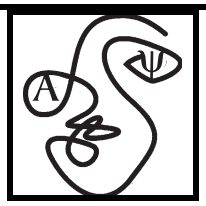
Meeus, W. H. J., & Raaijmakers, Q. A. W. (1987). Administrative obedience as a social phenomenon. *Current Issues in European Social Psychology*, *2*, 183–230.

Moghaddam, F., & Vuksanovic, V. (1990). Attitudes and behavior towards human rights across different contexts: The role of right-wing authoritarianism, political ideology and religiosity. *International Journal of Psychology*, *25*, 455–474.

Renteln, A. D. (1990). *International human rights: Universalism versus relativism*. London: Sage.

Shweder, R., Minow, M., & Markus, H. R. (2002). *Engaging cultural differences*. New York: Russell Sage Foundation.

United Nations Development Programme. (1996). *Human development report*. New York: Author.



Humor and Well-Being

James Rotton
Florida International University, University Park, Florida, USA

1. Multiple Meanings of Humor
2. Mechanisms: Possible Mediators and Moderators
3. Methodological Issues
4. Trait Humor, Distress, and Longevity
5. Humor and Physical Health
6. Conclusions
Further Reading

GLOSSARY

humor A broad and multifaceted concept that is variously used to refer to the quality of communications, the emotions that such communications elicit, and a tendency to respond with amusement in a variety of situations.

state humor Emotions (e.g., amusement, mirth) that link stimulus materials (e.g., comedy tapes) and observable responses (e.g., laughter).

trait humor A readiness to find or create humor in different situations (i.e., sense of humor).

well-being Physical and emotional health.

Articles on humor and health frequently begin by reminding readers, for example, that "a merry heart doeth good like a medicine" (from Proverb 17:22 in the King James Bible). Although few would argue with the first part of this proverb—"A merry heart doeth good"—there is very little agreement about whether or not humor should be likened to medicine. On the one hand, it is easy to come away from movies such as *Patch Adams* and articles in *Reader's Digest* with the impression that "laughter is the best medicine" (a phrase that has been used in the titles of several scientific articles). On the other hand, scientists doing research on health and well-being have expressed a considerable amount of skepticism about the medicinal effects of humor. Probably no article has done more to stimulate interest in humor and well-being than one that appeared in the *New England Journal of Medicine* by Norman Cousins, who was then the well-connected editor of the *Saturday Review*. In this brief article, which was subsequently expanded into a best-selling book, Cousins described how he checked himself out of a hospital after being diagnosed with a progressive, and usually fatal, form of rheumatoid disease. Following a self-prescribed regimen of Marx Brothers movies and *Candid Camera* television episodes, Cousins concluded that "10 minutes of genuine belly laughter had an anesthetic effect and would give me at least two hours of pain-free sleep."

1. MULTIPLE MEANINGS OF HUMOR

Humor is a broad and multifaceted concept that is variously used to refer to the quality of communications (e.g., funny or amusing stories), the emotions that such communications elicit (e.g., amusement, mirth), and a tendency to respond with amusement in a variety of situations (i.e., a sense of humor). As a stimulus, humor has been used to describe the

©2004 Elsevier Inc.
All rights reserved.

contents of literary material, television shows, movies, and (more recently) videotapes that are supposed to be funny. However, more than one study has reported checks on manipulations indicating that the individuals exposed to supposedly humorous material do not share the experimenters' definitions of "funny" or "amusing." This finding highlights the importance of individual differences. Few things are as idiosyncratic as people's preferences for different types of humor. If everybody agreed about what is funny or humorous, one would never have to endure the agony of seeing a comic "die on stage."

Serious problems are also encountered when one tries to define humor as a response or behavior such as laughing or smiling. No single behavior is a sure sign of amusement or mirth. For example, a subordinate may smile at the boss's joke for reasons that have nothing to do with how funny the joke is. To take another example, few movies provoke as much laughter in a college classroom as do Milgram's *Obedience to Authority,* which shows individuals forced to make stressful decisions. In such situations, laughter communicates anxiety rather than amusement. Furthermore, Provine's observational studies indicate that laughter in natural settings is typically preceded by mundane and pedestrian remarks. Provine could have used the familiar form of his first name, Bob, to illustrate this point. It was not long ago that the mundane greeting "Hi, Bob" (said by every person in a room) provoked laughter, first on *The Bob Newhart Show* and later whenever an individual named Robert entered a room.

Obviously, stimulus-based definitions of humor require that we observe some kind of response, and response-based definitions presume that evoking stimuli possess certain characteristics. These facts have led most theorists and researchers to define humor as an intervening variable. Following the lead of investigators doing research on emotions such as anxiety and anger, scholars frequently draw a distinction between state humor and trait humor. State humor describes emotions (e.g., amusement, mirth) that link stimulus materials (e.g., comedy tapes) and observable responses (e.g., laughter). Trait humor is a readiness to find or create humor in different situations. The latter corresponds to what is usually termed a person's "sense of humor." Investigators doing research on trait humor also draw a distinction between "getting a joke" (i.e., appreciative humor) and "telling a joke" (i.e., productive humor). In 1998, Craik and Ware identified several other types of humor. For example, the Humor Style Questionnaire highlights the well-known distinction between aggressive humor and friendly humor. Although the questionnaire's originators have been concerned primarily with people's health, research on aggression highlights the importance of drawing such a distinction. It has consistently been found that friendly humor reduces aggression, whereas witticisms (i.e., jokes at the expense of others) and hostile humor usually provoke hostility and/or aggression.

2. MECHANISMS: POSSIBLE MEDIATORS AND MODERATORS

At a time when individuals are constantly reminded that good health requires that they engage in activities they would rather avoid, such as exercising and eschewing sweet and fatty foods, it is not surprising that the public has seized on the idea that humor is good "for whatever may ail a person." Those who favor humor might not state things so bluntly (e.g., "Here's something you will like"), but they make a point of noting that there is a fairly strong connection between negative emotions and illness. It has been argued that if things that people do not like cause stress, which is associated with illness, then things that people like should add to health and longevity. Of course, the same argument could be made for sugar, fatty foods, and addictive substances.

Positive affect is usually assessed by having individuals complete self-report measures, although investigators occasionally obtain nonverbal measures such as frequency of laughter and smiling. These responses have been treated as separate dependent variables in research on humor and health. However, the case for mediation would be stronger if investigators performed analyses showing that relationships between humor and well-being vanished or were reduced after investigators statistically controlled for the intervening variable of positive affect. Unfortunately, such analyses have not been performed. In addition, there is no evidence to support the frequently heard claim that endorphins are released when individuals are exposed to humorous material.

Early research tried to identify cardiovascular and endocrinological changes that might be responsible for humor's effects. The physiological effects of laughter are similar to those of brief bouts of exercise—faster breathing, rapid heart rate, elevated skin temperature, and increased blood pressure. Higher levels of circulating catecholamines and cortisol have also been recorded. Early researchers suggested that people benefit from

the relaxation that follows laughter. However, these physiological changes have not been shown to be intervening variables that lead to long-term improvements in health and well-being.

Physiological explanations can be contrasted with cognitive approaches that emphasize appraisal and distraction. It is possible that humor enables individuals to cope with stress by helping them to view events as less threatening (i.e., cognitive appraisal) and by putting their problems into perspective (i.e., cognitive reappraisal). The latter is typified by jokes that begin, "You think you have problems." Another possibility is that the task of processing humor (i.e., getting a joke) distracts attention away from unpleasant events and experiences such as pain.

Finally, some of the effects attributed to humor may be due to the demonstrable effects of social support. In 2000, Provine pointed out that individuals rarely laugh when they are alone. Unfortunately, social support has been introduced as a confound by the common practice of testing individuals in small groups. Research is needed to determine whether humor exerts salubrious effects when individuals are tested alone.

3. METHODOLOGICAL ISSUES

The distinction between state humor and trait humor is relevant to the procedures that investigators employ to assess the effects of humor. By and large, researchers interested in the effects of state humor have adopted an experimental approach, whereas those conducting studies of trait humor have employed correlational techniques. Typically, in experiments on state humor, behavioral and physiological measures are collected after some participants have been exposed to an audio or video recording of a comic (nearly always Robin Williams or Bill Cosby). This has been done in both field and laboratory settings. It is difficult to interpret the results from laboratory studies when the dependent variable is pain tolerance because it is likely that most participants (especially undergraduate males) guess that the experimenters want to find out whether they can "grin and bear it." Regardless of whether studies are conducted in laboratory or field settings, researchers must take care to assess and control for their expectations and demand characteristics. So-called placebo effects are a problem because it is hard to disguise manipulations of humor. Because so much has been written about humor's supposedly beneficial effects, it is likely that most participants guess that the experimenter expects them to respond in a positive manner. This suggests that verbal ratings of stimulus materials, which are the usual checks on manipulation, might not be particularly useful or informative. As Provine pointed out, it is essential that researchers verify their inductions by monitoring behaviors such as smiling and laughter. Another necessity is to include a control group whose members are exposed to equally interesting and arousing but nonhumorous material. This type of control is needed to ensure that effects attributed to humor are not in fact caused by arousal or by distracting attention away from pain and unpleasant events.

In correlational investigations, participants typically complete several measures of humor and emotional states such as anxiety and depression. Two of the most widely used measures in this area are the Situational Humor Response Questionnaire (SHRQ) and the Coping Humor Scale (CHS). Lefcourt and Martin developed these to measure productive humor, but it is not entirely clear what these scales actually measure. Psychometric investigations indicate that these and other humor scales load heavily on a factor that Köhler and Ruch termed cheerfulness. The scales also tap a factor that has been termed "low seriousness" (or frivolity) that appears to accompany the detachment and perspective taking that humor is supposed to promote. Köhler and Ruch also found that correlations between self-report and behavioral measures of humor were low. These findings suggest that results obtained in this area might be due to other personality traits such as optimism, creativity, extraversion, and lack of social sensitivity. This is another way of saying that studies on trait humor lack discriminant validity. In addition, nearly all studies have been cross-sectional. Very few investigators have undertaken longitudinal investigations, which would strengthen claims that humor and well-being are linked in a causal fashion.

Although the distinction between trait humor and state humor is useful for organizing the literature, a few studies have combined the two approaches by measuring individuals' sense of humor before they are exposed to humorous material. It was thought that individuals with a well-developed sense of humor might benefit more from humorous viewing matter than would those with a less developed sense of humor. This hypothesis has not received support, but it is possible that people who have a sense of humor (or a cheerful disposition) go out of their way to expose themselves to material and situations that make them smile and laugh.

4. TRAIT HUMOR, DISTRESS, AND LONGEVITY

As might be expected, measures of trait humor are correlated with scales that assess positive emotions such as happiness and subjective well-being. Conversely, they are negatively correlated with measures of distress (e.g., anxiety, depression, hostility). However, this pattern may reflect little more than the operation of a more general trait such as cheerfulness. Thus, it is noteworthy that early studies by Lefcourt and Martin found that trait humor may act as a buffer that reduces the negative effects of stressful life events. These investigators concluded that humor is "an antidote to stress" after they had undergraduates complete three sets of self-report measures, including the previously mentioned measures of humor (SHRQ or CHS) and a measure of stressful life events. They used scores on these two measures to predict scores on a third measure (total mood disturbance), which was the sum of four negative mood states (e.g., anxiety, depression) and a negative score for positive mood. They found that some, but not all, types of humor acted as a stress-reducing buffer. In particular, people who scored high on productive humor (i.e., attempts at being funny) were less likely to report negative mood states when they encountered stressful life events than were those who scored low on productive humor. However, this finding received mixed support in subsequent studies. Overall, productive humor does not appear to be any more effective than humor appreciation in helping people to cope with stress. In addition, the effects of trait humor appear to depend on how stress is defined (i.e., stronger effects when adversity is recalled than when it is experienced), on how distress is operationalized (i.e., stronger effects for depression than for anxiety), and on gender (i.e., stronger effects for males than for females). These inconsistencies led Martin and Lefcourt to use words such as "ambiguous" and "uncertain" to describe the buffering effects of trait humor.

Suggestive and sometimes intriguing findings have also been obtained in studies that have examined relationships between trait humor and physical symptoms. Unfortunately, all of the findings are correlational and based on self-report measures. Because investigators have not controlled for negative affectivity, it is possible that people with a sense of humor report fewer symptoms because they are the kind of people who are less likely to appraise events in a negative fashion.

Nonsignificant and sometimes contrary results have been obtained when investigators have tried to relate trait humor to more objective measures such as longevity. One study found that famous humorists and comics, who earned their reputations by producing humor, did not live any longer than individuals who achieved eminence in other fields. Granted, this study can be faulted for failing to draw a distinction between friendly humorists (e.g., Fred Allen, Will Rogers) and comics whose witticisms and insults are associated with negative emotions (e.g., Don Rickles). This criticism does not apply to a longitudinal analysis of mortality rates of gifted individuals who were first tested at 12 years of age as part of Terman's Life Cycle Study. Friedman and colleagues combined teacher ratings of optimism and sense of humor into a measure of cheerfulness. Contrary to expectations, they found that individuals with a cheerful disposition died at earlier ages than did their more serious counterparts. This unexpected result was explained by suggesting that a cheerful disposition might lead individuals to ignore or discount threats to their health.

5. HUMOR AND PHYSICAL HEALTH

In addition to longevity, investigators have looked at the effects of humor on blood pressure, immune function, and pain tolerance. It appears that humor does not have long-term effects on cardiovascular measures such as heart rate and blood pressure. However, there is some reason to believe that trait humor acts as a buffer to reduce the systolic blood pressure of females but not that of males. It may be that females use humor to reappraise threatening events (e.g., "Look on the bright side"), whereas males show a preference for aggressive humor.

Mixed results have been obtained in studies that have looked at the effects of humor on the body's immune system. On the one hand, a few positive results have been obtained in studies that relied on blood assays, but immunological indicators vary from one study to the next—sometimes natural killer (NK) cell activity, sometimes interleukin cell activity, sometimes both, but more often neither. It is difficult to interpret these findings because investigators have not controlled for general emotional arousal and the pyramiding of Type I errors that occurs when a large number of tests of significance are conducted. On the other hand, several studies have found that humor leads to higher levels of secretory immunological A (SIgA), which is the body's first line

of defense against upper respiratory illnesses. However, nonsignificant results were obtained when investigators looked at the frequency of cold and flu symptoms. Of greater concern, studies on SIgA have lacked adequate controls for general emotional arousal, distraction, and diurnal variations in hormone production and immune function.

Finally, although research on most health variables has been disappointing, it is in some ways fitting that studies on pain provide partial vindication for the claims that Cousins originally made in 1976. Laboratory studies are fairly consistent in finding that individuals have a higher tolerance for pain after watching an amusing videotape than they do after watching one that is merely interesting or distracting. There is also some evidence that humorous materials help individuals to endure orthopedic and arthritic pain in hospital settings. For example, in a field experiment in 1996, Rotton and Shats compared the amount of pain medication that patients requested after watching either comedies or adventure movies following orthopedic surgery. They found that patients who watched comedies requested smaller doses of minor analgesics (e.g., aspirin) than did those who watched adventure movies or who were not given movies to watch. However, this beneficial effect did not extend to major analgesics (e.g., Percodan). Indeed, the results indicated that patients actually asked for more medication when they were not given a choice as to which films they could watch. As the researchers observed, having no choice but to watch comedies that fail in their attempts to be amusing probably elicits negative emotions.

6. CONCLUSIONS

Belief in the healing power of humor is so widespread that this article's skeptical tone might seem surprising to some. Those who question humor's medicinal effects run the risk of being labeled curmudgeons. Thus, it should be noted that this article's evaluation echoes conclusions found in books and articles by the most active and frequently cited researchers in this field. It should be cautioned that these conclusions are based on a relatively small number of empirical studies. For example, Martin reported that his recent review of health effects uncovered no more than 45 empirical studies.

It is possible to offer two reasons why very little research has been done to untangle relationships between humor and well-being. First, belief in the healing power of humor is so widespread that it already seems to be taken as proven. That is, why do research to test a hypothesis that most people already believe is true? Second, one may recall that the concept of humor appears to load onto two factors: cheerfulness and lack of seriousness. Most people do not object to being termed "cheerful," but they find it difficult to have their work dismissed as lacking in seriousness. These two tendencies may explain why publications in this area rarely include footnotes that acknowledge funding sources.

Reviews usually address a paucity of research by suggesting that more is necessary. However, what is really needed is better research—experiments that contain necessary control groups and verifiable measures of humor (e.g., laughter, smiling), studies that rule out general arousal as a rival hypothesis, intervention studies that avoid all of the biases (e.g., expectancies, placebo effects) that self-report measures introduce, investigations whose results are based on a sufficient number of participants, and research that explores the possible mediating role of social support.

Rather than comparing humor to medicine, as has been done so frequently, it might be more appropriate to compare the effects of humor to those that may be produced by vitamin pills and the supplements one finds at health food stores. The government prohibits sellers from claiming that dietary supplements cure anything, but those who buy vitamins believe that they benefit from ingesting them. In the same way, the literature in this area contains intriguing results from unfortunately flawed studies suggesting that individuals may sometimes benefit from a dose of laughter. Stores can sell vitamins because they do not contradict the physician's edict: "First do no harm." The same can be said of humor. There is probably no harm in a regimen that keeps people in stitches—unless, of course, an individual is recovering from an appendectomy.

See Also the Following Articles

Emotion ■ Emotion and Culture ■ Social Comparison and Subjective Well-Being ■ Stress ■ Traits ■ Well-Being

Further Reading

Cousins, N. (1976). Anatomy of an illness (as perceived by the patient). *New England Journal of Medicine*, *293*, 1458–1463.

Craik, K. H., & Ware, A. P. (1998). Humor and personality in everyday life. In W. Ruch (Ed.), *The sense of humor* (pp. 63–94). New York: Guyton de Gruyter.

Köhler, G., & Ruch, W. (1998). Sources of variance in current sense of humor inventories: How much substance, how much method variance? *Humor*, *9*, 305–321.

Lefcourt, H. M. (2002). Humor. In C. R. Snyder, & S. J. Lopez (Eds.), *Handbook of positive psychology* (pp. 619–631). New York: Oxford University Press.

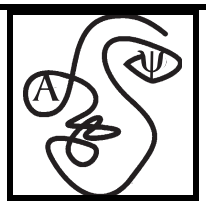
Martin, L. R., Friedman, H. S., Tucker, J. S., Tomlinson-Keasey, C., Criqui, M. H., & Schwartz, J. E. (2002). A life course perspective on childhood cheerfulness and its relation to mortality risk. *Personality and Social Psychology Bulletin*, *28*, 1155–1165.

Martin, R. A. (2003). Sense of humor. In S. J. Lopez, & C. R. Snyder (Eds.), *Positive psychological assessment: A handbook of models and methods* (pp. 313–326). Washington, DC: American Psychological Association.

Martin, R. A., Puhlik-Doris, P., Larson, G., Gray, J., & Weir, K. (2003). Individual differences in uses of humor and their relation to psychological well-being: Development of the Humor Styles Questionnaire. *Journal of Research in Personality*, *37*, 48–75.

Provine, R. R. (2000). *Laughter: A scientific investigation*. New York: Penguin.

Rotton, J., & Shats, M. (1996). Effects of state humor, expectancies, and choice on postsurgical mood and self-medication: A field experiment. *Journal of Applied Social Psychology*, *26*, 1775–1794.



Hypnosis

John F. Kihlstrom
University of California, Berkeley, California, USA

1. Hypnotizability and Its Correlates
2. Pain and Analgesia
3. Other Clinical Applications
4. Sensorimotor Performance Enhancement
5. Learning and Memory
6. Risks of Hypnosis
7. Conclusion
   Further Reading

GLOSSARY

age regression In hypnosis, an experience in which the hypnotized person returns to a scene or an event experienced in the past.

analgesia A reduction in pain.

anesthesia A reduction in tactile sensation; in surgery, this is often accompanied by a general loss of consciousness.

conscious sedation A surgical technique in which the patient remains conscious during a procedure but receives analgesic drugs to reduce pain as well as sedative drugs to reduce anxiety and induce amnesia for the procedure.

false alarms In memory testing, an instance where the person incorrectly identifies an item as having appeared on a previously studied list.

hemispheric specialization The idea that the right and left hemispheres of the cerebral cortex are specialized for different psychological functions.

hyperesthesia Enhancement of sensory acuity and other perceptual abilities; it is the opposite of anesthesia.

hypermnesia Enhancement of memory; it is the opposite of amnesia.

hyperpraxia Enhancement of muscular strength, endurance, and motor ability; it is the opposite of apraxia.

hypnotizability Individual differences in response to hypnosis, as measured by standardized psychological tests such as the Stanford Hypnotic Susceptibility Scales.

openness to experience A "Big Five" personality trait characterized by richness in fantasy life, aesthetic sensitivity, awareness of inner feelings, need for variety in actions, intellectual curiosity, and liberal values.

psychodynamic Broad label for a class of theories of personality and psychotherapy based on Freud's idea that behaviors are motivated by unconscious conflict.

psychosomatic The term applied to medical illnesses and other physical conditions that appear to be caused by mental states such as particular beliefs and attitudes.

self-fulfilling prophecy A situation in which a person's beliefs and expectations lead him or her to behave in such a manner as to make the expectations come true, thereby confirming the person's initial beliefs.

Hypnosis is a social interaction in which one person responds to suggestions given by another person (the hypnotist) for imaginative experiences involving alterations in perception, memory, and the voluntary control of action. In the classic instance, these suggested experiences are associated with a degree of subjective conviction bordering on delusion and an experience of involuntariness bordering on compulsion. Hypnosis has its origins in the "animal magnetism" promoted by Franz Anton Mesmer, an Austrian physician, for the treatment of medical illnesses.

©2004 Elsevier Inc.
All rights reserved.

Mesmer's theories were discredited in 1784 by a French royal commission chaired by Benjamin Franklin, but interest in his technique persisted. Over the years, animal magnetism was transformed into "mesmerism," then "artificial somnambulism," then "neurohypnotism," and finally hypnosis as we know it today. Since the time of William James and Ivan Pavlov, hypnosis has been of interest to psychologists and other scientists and philosophers for the insights it can yield about perception, memory, consciousness, and other basic psychological functions. However, it has also found its way into practical applications of various sorts.

1. HYPNOTIZABILITY AND ITS CORRELATES

It is important to understand that not everyone responds positively to hypnotic suggestions. Individual differences in hypnotizability are measured by performance-based tests such as the Stanford Hypnotic Susceptibility Scale and the Harvard Group Scale of Hypnotic Susceptibility. Some researchers and clinicians use alternative scales, such as the Hypnotic Induction Profile, but the Harvard scale is the most popular choice for this purpose and the Stanford scale remains the "gold standard" for measuring hypnotizability. The hypnotizability scales consist of a standardized induction procedure accompanied by a set of representative hypnotic suggestions, the response to which is scored according to objective behavioral criteria. As measured by the scales, most people are at least moderately responsive to hypnotic suggestions; relatively few people are refractory to hypnosis, and relatively few "hypnotic virtuosos" fall within the highest level of responsiveness. Some theorists have argued that hypnotizability can be enhanced by developing positive attitudes, motivations, and expectancies concerning hypnosis. As with any other skilled performance, hypnotic response is probably a matter of both aptitude and attitude. Negative attitudes, motivations, and expectancies can interfere with performance, but positive ones are not by themselves sufficient to create high hypnotizability.

Hypnotizability is not substantially correlated with most other individual differences in ability or personality such as intelligence and extraversion. One exception is absorption, that is, the tendency to have subjective experiences characterized by the full engagement of attention (narrowed or expanded) and blurred boundaries between self and object. Absorption, in turn, is a facet of the "openness to experience" dimension in the "Big Five" structure of personality. However, even these correlations are too weak to permit hypnotizability to be predicted in advance on the basis of the usual sorts of personality questionnaires.

Although hypnosis is commonly induced by suggestions for relaxation and even sleep, the brain activity in hypnosis more closely resembles that of a person who is awake. The discovery of hemispheric specialization led to the speculation that hypnosis was a "right-brain" function, but it is obvious that response to hypnotic suggestions also requires the verbal functions normally associated with the left cerebral hemisphere. Because the experience of involuntariness is a central feature of hypnosis, some theorists have proposed that the state entails alterations in functioning of the frontal lobes of the brain that are known to be involved in executive control. However, a better understanding of the neural substrates of hypnosis awaits studies of neurological patients with focalized brain lesions as well as brain-imaging studies (e.g., positron emission tomography [PET], functional magnetic resonance imaging [fMRI]) of normal individuals.

2. PAIN AND ANALGESIA

By far, the most successful clinical application of hypnosis involves suggestions for analgesia to relieve pain. As early as the 1840s, John Elliotson and James Esdaile each reported that they had performed a large number of surgeries painlessly with artificial somnambulism as the only anesthetic agent. However, beginning in 1846, the introduction of chemical anesthetics, such as ether and chloroform, quickly supplanted hypnosis. Nevertheless, when chemical analgesics or anesthetics were unavailable or contraindicated, physicians still sometimes turned to hypnosis. Between 1955 and 1974, more than two dozen cases in which hypnosis had been used as the sole analgesic or anesthetic agent in surgery were published. Of even greater importance was the publication, during the early 1960s, of extensive case series documenting the successful use of hypnosis in obstetrics and in the treatment of cancer pain. Other clinical studies found that hypnosis can effectively relieve pain in dental patients. One comprehensive review of the clinical literature estimated that approximately half of medical, dental, and surgical patients could obtain significant pain relief through hypnosis alone. More recent reviews confirm that even when patients are unselected for

hypnotizability, hypnotic suggestions can have quite substantial effects on both acute and chronic pain. Hypnosis may be especially useful in cases of chronic pain, where chemical analgesics such as morphine pose risks of tolerance and addiction, and in childbirth, where the active participation of the mother may be desirable.

In addition to replacing chemical analgesics when circumstances warrant, hypnosis can serve as a useful adjunct to more conventional medical treatment of pain such as "conscious sedation" for outpatient surgery. Careful clinical studies by at least three independent groups of investigators have shown that the addition of hypnosis reduces both pain and anxiety. Patients request, and receive, less pain medication than do controls. Hypnosis is associated with a decrease in interruptions and adverse events during surgery. The procedure takes less time, and patients recover faster, when patients receive both hypnosis and medication. Although hypnosis can rarely substitute for the chemical analgesia and anesthesia that is the standard of care in modern medicine, it is a cost-effective adjunctive treatment (a complement if not an alternative) that improves the quality of patient care.

In addition to this increasing body of clinical research, a large amount of laboratory research has shed light on the mechanisms by which hypnotic suggestions have their effects. Hypnotic analgesia is not mediated by mere relaxation or by the release of endogenous opiates (endorphins). It is not merely a placebo effect (although there is a placebo component in hypnotic analgesia, just as there is in chemical analgesics). Hypnosis can reduce both sensory pain and suffering; in fact, suggestions specifically targeting one or the other of these components of pain have selective effects on activity in the somatosensory and anterior cingulate regions of the brain. Psychologically speaking, hypnotic analgesia seems to go beyond self-distraction, stress inoculation, cognitive reinterpretation, and tension management. According to one prominent theory, hypnotic analgesia involves a dissociation, or a division of consciousness, that effectively reduces patients' awareness of pain.

3. OTHER CLINICAL APPLICATIONS

Hypnotic suggestion can have "psychosomatic" effects on aspects of bodily functioning other than pain, and these can also be useful in the practical sense. For example, it has frequently been observed that hypnotic suggestions can reduce bleeding from wounds and speed the healing of burns. However, these effects have not yet been subject to rigorous clinical and experimental study. Several well-controlled laboratory and clinical studies have shown that hypnotic suggestion can affect allergic responses, asthma, and the remission of warts. Such successes have led some practitioners to offer hypnosis in the treatment of cancer, but there is no evidence that hypnosis affects tumors themselves as opposed to patients' quality of life.

Hypnosis has also been used in psychotherapy. Some psychodynamic theorists, working in the tradition of Sigmund Freud, consider hypnosis to be a form of "adaptive regression" or "regression in the service of the ego," and they use it to promote relaxation, enhance imagery, and generally loosen the flow of free associations. However, there is little evidence from controlled outcome studies that "hypnoanalysis" or "hypnotherapy" is more effective than nonhypnotic forms of the same treatment. In contrast, a recent review found substantial treatment gains when hypnosis was used adjunctively in cognitive–behavioral therapy for problems such as obesity, insomnia, anxiety, phobia, and hypertension. In view of the demands of managed medical and mental health care, it is important for practitioners who use hypnosis to document quantitatively the clinical (and economic) benefits of doing so.

An important but unresolved issue is the role played by individual differences in the clinical effectiveness of hypnosis. In the laboratory, as in the clinic, a genuine effect of hypnosis should be correlated with hypnotizability. Unfortunately, clinical practitioners are often reluctant to assess hypnotizability in their patients and clients due to a concern that low scores might reduce motivation for treatment. This danger is probably exaggerated. On the contrary, assessment of hypnotizability by clinicians who are contemplating the therapeutic use of hypnosis would seem to be no different, in principle, from assessment of allergic responses before prescribing an antibiotic. In both cases, the legitimate goal is to determine what treatment is appropriate for what patient.

It should be noted that clinicians sometimes use hypnosis in nonhypnotic ways, and these practices tend to support the hypothesis that whatever effects they achieve through hypnosis are related to its placebo component. For example, there is nothing particularly "hypnotic" about having a patient in a smoking cessation treatment rehearse therapeutic injunctions not to smoke and other coping strategies while hypnotized. It is possible that many clinical benefits of hypnosis are mediated by placebo-like motivational and expectational processes, that is, by the "ceremony" surrounding hypnosis rather than by hypnosis per se.

4. SENSORIMOTOR PERFORMANCE ENHANCEMENT

Much of the popular interest in hypnosis stems from claims that it permits people to transcend their normal voluntary capacities, yielding increased muscular power, endurance, resistance to fatigue, and sensory acuity as well as other special abilities. However, controlled research has largely failed to find evidence that hypnosis can enhance these and other aspects of human performance. Many early studies, which seemed to yield positive results for hypnosis, possessed serious methodological flaws such as the failure to collect adequate baseline information. For example, individuals may "hold back" on nonhypnotic trials so as to satisfy experimenters' implicit expectations that they show "improved" performance during hypnosis. Of course, in some situations, benefits may be achieved indirectly through suggestions for reduced awareness of pain. For example, hypnotized persons may experience themselves as stronger, but this is not the same as a direct enhancement of human performance capacities per se.

In general, it appears that hypnotic suggestions for increased muscular strength, endurance, sensory acuity, and/or learning do not exceed what can be accomplished by appropriately motivated persons outside of hypnosis. Nevertheless, the folklore surrounding hypnosis may lead hypnotized persons to achieve performance enhancements through a kind of self-fulfilling prophecy. For example, instead of giving suggestions for increased muscle strength, one intriguing study asked hypnotizable persons to imagine themselves becoming stronger. These individuals actually outperformed controls who were merely exhorted to maximize their performance. If hypnosis does enhance human performance, it may be by virtue of hypnotic alterations in individuals' self-concepts rather than by virtue of their skeletal musculatures.

Another highly intriguing line of research contends that hypnotic suggestions may enhance visual acuity in individuals with myopia (although not in individuals with normal vision). The research is especially intriguing because the investigators took special care to rule out "peripheral" factors, such as accommodation and other structural changes in the eyes, that might account for the effect. If hypnotic suggestions do in fact improve visual acuity in people with impaired vision (and this claim still must be confirmed by additional research), the effects appear to be mediated centrally, perhaps involving cortical centers associated with vision or visual attention.

It is sometimes claimed that hypnosis enhances "parapsychological" abilities such as clairvoyance, telepathy, and telekinesis. However, there is no good evidence that this is the case, not least because there is no good evidence of parapsychological effects in the first place.

5. LEARNING AND MEMORY

Similar conclusions apply to learning capacity. Many studies of hypnotically enhanced learning suffer from the same methodological defects as do the aforementioned studies of muscular performance and sensory acuity. One well-known line of research attempted to enhance learning by giving individuals hypnotic suggestions for time distortion. By extending the "subjective" interval of time devoted to studying the material, these investigators hoped to enable these individuals to learn more over short intervals of time. However, the promising results of early experiments have largely failed to be confirmed by more tightly controlled follow-up studies. Hypnotized persons may well experience time as flowing more slowly, or more rapidly, than it actually is, but this distortion in perception does not have positive consequences for learning.

A special case of performance enhancement pertains to hypnotic suggestions for improvements in memory or "hypnotic hypermnesia." Hypermnesia suggestions are sometimes employed in forensic situations, to help forgetful witnesses and victims remember, and in therapeutic situations, to help patients remember childhood sexual abuse and other traumatic personal experiences. Sometimes, suggestions for enhanced memory are accompanied by suggestions for age regression in which hypnotized persons are asked to return to a previous period in their lives. Although the clinical literature contains a number of reports claiming success, most of these are anecdotal in nature and fail to obtain independent corroboration of the memories that emerge. There is no clinical evidence that hypnotically enhanced memories are reliable, that is, that they can be taken as true without independent confirmation.

In fact, evidence from carefully controlled laboratory studies shows quite the opposite. Hypnosis does not appear to increase accurate recollection over and above what can be achieved under nonhypnotic circumstances. To make things worse, hypnosis appears to increase the likelihood of "false alarms" on recognition tests and to inflate individuals' confidence in their memories, independent of their accuracy. Moreover, by virtue of their

enhanced responsiveness to suggestion, hypnotized persons may be especially vulnerable to leading questions and other suggestive influences that might bias or distort their memories. Hypnotized persons may well believe that they are recovering forgotten memories, but that does not make the memories valid representations of the historical past. Any memories recovered by means of hypnosis must be subject to independent corroboration before they are taken seriously.

Similar considerations apply to hypnotic age regression. Although age-regressed individuals may experience themselves as children and may behave in a distinctly childlike manner, there is no evidence that they actually undergo either abolition of characteristically adult modes of mental functioning or reinstatement of childlike modes of mental functioning. Nor do age-regressed individuals experience the revivification of forgotten memories of childhood.

Given what we know about the unreliability of hypnotic hypermnesia, the use of hypnosis to recover memories in clinical and forensic situations must be judged to be highly risky and not recommended. In fact, many legal jurisdictions severely limit the introduction of memories recovered through hypnosis due to a concern that such evidence might be tainted. The Federal Bureau of Investigation (FBI) has published a set of guidelines for those who wish to use hypnosis forensically, and similar precautions should be employed in the clinic. Although field studies have sometimes claimed that hypnosis can enhance memory in a powerful way, these anecdotal reports have not been duplicated under laboratory conditions.

6. RISKS OF HYPNOSIS

Hypnosis poses few risks to persons who are hypnotized. Because it requires the active participation of the hypnotized persons, people cannot be hypnotized unknowingly or against their will. Nor can even deeply hypnotized persons be led, by virtue of suggestion, to do something that they would not do under other circumstances. In experimental situations, the risks to individuals are usually limited to drowsiness or mild headache—no different in kind or severity from those that crop up when students take exams or listen to lectures. In most instances, these transient problems can be rapidly traced to some specific expectation on the part of the hypnotized persons, perhaps derived from inadvertent comments made by a stage performer or television commentator seen recently. Occasionally, a hypnotized person actually dozes off during the procedure or rests his or her head in an uncomfortable position, leading to drowsiness or neckache on arousal. In any event, a few minutes of supportive interaction are usually sufficient to deal with the problem effectively. However, problems may be compounded when inexperienced experimenters find it difficult to cope with a minor problem and communicate their lack of competence and self-confidence to the hypnotized person, who may then proceed to make matters even worse. Problems may also occur in stage hypnosis, where the entertainment context has special effects on experience and behavior over and above the hypnotic context.

Untoward consequences are more likely to arise when hypnosis is employed during therapeutic, as opposed to experimental, procedures. A manipulative therapist who simply tries to "suggest away" symptoms may encounter problems that stem from his or her own therapeutic errors. When hypnosis is used to recover ostensibly forgotten memories, whether in therapeutic or forensic situations, the memories of the hypnotized person may be contaminated with information that is objectively false—with potentially serious negative consequences for other people. When hypnosis is used to control pain in athletic situations, competitors may ignore important information about the location and severity of bodily damage. In other cases of performance enhancement, it is possible that individuals will harm themselves through overexertion. Practitioners should not attempt to treat with hypnosis any condition that they are not qualified to treat without hypnosis.

7. CONCLUSION

Claims of hypnotically induced performance enhancement are nearly as old as hypnosis itself. In George DuMaurier's novel *Trilby,* published in 1895, Svengali transforms the eponymous heroine into an accomplished singer. And reviewing the hypnosis literature nearly half a century ago, F. L. Marcuse referred to "The Generation of Hypers," that is, the notion that hypnosis could produce hyperpraxia, hyperesthesia, and hypermnesia as well as other special effects. Aside from the ability of hypnotic suggestion to control pain—a genuine phenomenon of substantial practical importance—most claims for hypnotic performance enhancement are dubious at best. Hypnosis is fascinating to observe, is interesting from a theoretical standpoint, and has many actual and potential

applications as an adjunctive treatment in medicine and psychotherapy, but it will not make us better people.

Acknowledgments

The point of view represented in this article is based on research supported by Grant MH-35856 from the National Institute of Mental Health.

See Also the Following Articles

Learning ■ Pain Management ■ Psychoneuroimmunology ■ Psychotherapy in Older Adults

Further Reading

Bowers, K. S. (1976). *Hypnosis for the seriously curious*. Pacific Grove, CA: Brooks/Cole.

Fromm, E., & Nash, M. R. (Eds.). (1992). *Contemporary hypnosis research*. New York: Guilford.

Gauld, A. (1992). *A history of hypnotism*. Cambridge, UK: Cambridge University Press.

Hilgard, E. R. (1965). *Hypnotic susceptibility*. New York: Harcourt, Brace, & World.

Hilgard, E. R., & Hilgard, J. R. (1975). *Hypnosis in the relief of pain*. Los Altos, CA: Kaufman.

Kihlstrom, J. F., & Barnhardt, T. M. (1993). The self-regulation of memory: For better and for worse, with and without hypnosis. In D. Wegner, & J. W. Pennebaker (Eds.), *Handbook of mental control* (pp. 88–125). Englewood Cliffs, NJ: Prentice Hall.

Kihlstrom, J. F., & Eich, E. (1994). Altering states of consciousness. In D. Druckman, & R. A. Bjork (Eds.), *Learning, remembering, believing: Enhancing human performance* (pp. 207–248). Washington, DC: National Academy Press.

Kirsch, I., Montgomery, G., & Sapirstein, G. (1995). Hypnosis as an adjunct to cognitive-behavioral psychotherapy: A meta-analysis. *Journal of Consulting and Clinical Psychology*, *63*, 214–220.

Nash, M. R. (2001, July). The truth and the hype of hypnosis. *Scientific American,* pp. 47–55.

Patterson, D. R., & Jensen, M. P. (2003). Hypnosis and clinical pain. *Psychological Bulletin*, *129*, 495–521.

Ideological Orientation and Values

Hans De Witte
Catholic University of Leuven, Leuven, Belgium

1. Ideology and Values: Definitions
2. Values and Value Orientations in Political Psychology
3. On the Dimensionality of (Conservative) Ideology
4. Political Consequences of Ideology and Value Orientations
Further Reading

GLOSSARY

attitude A relatively enduring organization of beliefs around a concrete object or situation, predisposing people to respond in some preferential manner.

competition Value orientation that combines the acceptance of inequality with a focus on the individual. Central values are commitment to achieving wealth, social recognition, and social power.

conformity **(to external authorities)** Values stressing obedience to external norms and the need to follow the dictates of authority.

ideology A coherent set of attitudes and values, shared by a social group, that motivates their actions.

materialist Values expressing the need for economic and financial security, societal stability, personal safety, and law and order.

post-materialist Values referring to the need for political freedom and participation, self-actualization, personal relationships, creativity, and care for the environment.

self-direction Values emphasizing autonomy and the personal responsibility of the individual.

(socio-)cultural conservatism versus progressiveness Also labeled "sociocultural libertarianism–authoritarianism." Ideological opposition related to the old cleavage between the Church and the State, typical of pre-industrial sources of political conflict. Conservatism emphasizes traditional norms (e.g., traditional gender roles and strict child rearing practices). Progressiveness puts traditional norms into perspective, stressing, e.g., women's emancipation and the autonomy of children, and supporting the liberalization of abortion.

(socio-)economic conservatism versus progressiveness Also labeled "socio-economic left-right." Ideological opposition related to socioeconomic cleavages in industrial society. A progressive stand rejects social inequality as unfair and argues for government intervention and stronger trade unions in order to reduce economic inequality. Conservatives oppose government intervention and (militant) trade unions, and emphasize private initiative and individual competition as essential for proper economic functioning.

universalism Value orientation that combines equality and focus on others into a universalistic, altruistic value pattern. Equality, social justice, and true friendship are key values in this orientation.

values Abstract, fundamental goals. They are evaluative standards with normative power. Values are hierarchically ordered in a coherent pattern.

This article discusses the content and relevance of ideology and values for political psychology. Even though certain scholars advocated "the end of ideology" in the past, recent events suggest that ideology still plays an important role in contemporary society, in shaping and affecting the life of everyday people and of specific, politically active groups. The

©2004 Elsevier Inc.
All rights reserved.

terrorist attacks of September 11, 2001 are a powerful example of the impact of an extreme ideology on innocent citizens. This event also demonstrates that radical ideologies can motivate activists to perform extreme acts, by sacrificing their life for a "higher cause." Less drastic examples of the relevance of ideology can be found in the rise of new political parties on the left (e.g., green parties) and the right (e.g., extreme right-wing parties), which attracted more voters than before during parts of the last decade.

In this article, first, the concepts of ideology and values are defined. Next, influential theories on values and value orientations in political psychology are reviewed, starting with the less complex ones, which will be integrated in more recent multidimensional theories at the end. The third section discusses the dimensionality of conservatism, and relates it to the values section, and to the specific groups that are holding this ideology. Fourth, the political consequences of ideology and value orientations are outlined. The article concludes with some crucial references for further reading.

1. IDEOLOGY AND VALUES: DEFINITIONS

The term ideology is used in various disciplines such as psychology, sociology, and political science. Consequently, it has many meanings. In psychological research, however, consensus exists regarding three core aspects of an ideology: it is a coherent set of attitudes and values, it is shared by and therefore typical of a given social group, and it motivates their actions.

Ideology is not a mixed bag of attitudes. Rather, it is a logical and (relatively) coherent world view and "portrayal of man" that provides an understanding of reality: it offers an interpretation and an explanation of how the world works and of the role and motives of the people in that world. This world view and vision of mankind are founded on certain values. As a consequence, an ideology does not only describe "what is" (first aspect), but also "what ought to be." So, ideologies are not purely descriptive, they are also prescriptive or normative: they indicate what should be achieved and changed. They describe the desired goal and direct or motivate action (third aspect). This mobilizing function, or incentive to act, is linked to the collective nature of an ideology: ideologies refer to a set of attitudes and values that are shared by a larger group (second aspect). They reflect the interests of a specific social group and offer identification possibilities to the members of these groups. By accepting the collective ideology, they can show that they too belong. These shared viewpoints can be used to make a distinction between those who belong ("us") and those to whom this does not apply ("them"). An ideology also describes the group goals that must be achieved in order to secure the group's interests or to improve the position of the group and its members.

Attitudes and values are part of an ideology. This means that these concepts must also be defined. A wide variety of definitions is used for the concept of "values" in the literature. There is, however, consensus on a number of basic features. Values are fundamental goals. They are evaluative standards with normative power: they control behaviors, thoughts, and feelings. Values are abstract, not directly observable constructs. They are not isolated from each other, but form a coherent pattern that is hierarchically ordered. Finally, values also have a certain continuity and stability. Most recent definitions of values broadly embrace these aspects of the concept. Schwartz (1994), for example, defines a value as "a belief pertaining to desirable end states or modes of conduct, that transcends specific situations, guides selection or evaluation of behavior, people and events, and is ordered by importance relative to other values to form a system of value priorities" (p. 20).

Rokeach, who can be considered one of the founding fathers of value research, emphasized the distinction between values and attitudes. He defined attitudes as a relatively enduring organization of beliefs around an object or situation, predisposing one to respond in some preferential manner. Attitudes are thus more concrete than values, because they refer to a specific object or specific situation. Values have a more abstract and fundamental character. A set of items that relates to the evaluation of income differences and to government intervention in the economy can therefore be considered as a set of attitudes. Because values are more fundamental, they are also more limited in number. An individual can subscribe to a large number of attitudes, but only to a limited number of values. Attitudes thus form "concretizations" of values, because a single fundamental value can be translated into a large number of attitudes with regard to actual topics. Finally, Rokeach also assumes that values occupy a more central position than attitudes within one's personality and cognitive system. As a consequence, they are determinants of attitudes and behaviors.

2. VALUES AND VALUE ORIENTATIONS IN POLITICAL PSYCHOLOGY

Research into values and value orientations is of course not restricted to political psychology. In this field, numerous contributions of scholars in other fields of the social sciences and psychology are used and applied to the political sphere. The following sections discuss some of the most influential, starting with the most limited ones, focusing on just a selection of values. These will be integrated in more recent multidimensional accounts later on.

2.1. Kohn: Conformity versus Self-Direction and Its Relationship with Social Class

Kohn and his co-workers performed important research on the relationship between values and the position of individuals in the social structure. Their research shows that the middle class is oriented toward the value of self-direction. Respondents from the middle class find it more desirable to behave on the basis of their own insights and emphasize self-direction, autonomy, and the personal responsibility of the individual. In short, they value thinking for oneself. This value orientation has important consequences for their behavior in a variety of life domains. One of these domains is the upbringing of children, in which they value aspects such as consideration, curiosity, and self-control. Blue-collar workers, on the other hand, emphasize conformity to external authorities: they stress obedience to external norms, and the need to follow the dictates of authority. Examples of values stressed in the upbringing of children are obedience, good manners, and neatness. These scholars also demonstrate that self-direction and conformity (to external authority) are opposite poles of a single continuum.

Kohn and his co-workers stress that the difference in attitudes and values between blue-collar workers and white-collar workers is caused by the extent to which their work requires "occupational self-direction." This variable refers to autonomy in the workplace and is composed of three components: the complexity of the work, the amount of supervision, and the (non-)repetitive character of the job. The extent to which someone can work autonomously depends on his or her occupational position. Employees in higher professions can work more autonomously: their work is more varied, less supervised, and more complex, so that they must make more decisions on their own. Through longitudinal research, Kohn and his co-workers demonstrated that one's capacity to work autonomously influences the values one develops. Work that requires autonomy leads to an orientation toward self-direction. When the work does not allow autonomy, then one has to adapt to the (work) environment by emphasizing conformity. The logic behind this link is that obeying external authorities forms the best adjustment to a situation in which one has little control of the course of events.

2.2. Inglehart: Materialism versus Post-Materialism

Inglehart's post-materialism thesis is an example of a theory that advocates the social and political relevance of a new set of values in society. According to Inglehart, advanced Western societies experience a shift from materialist to post-materialist values. Material values express the need for economic and financial security, societal stability, personal safety, and law and order. Post-materialist values refer to the need for political freedom and participation, self-actualization, personal relationships, creativity, and care for the environment. In modern societies, a "silent revolution" is taking place, in which the dominance of materialist values is being replaced by post-materialist values. This value change is caused by the continuous replacement of old generations by new ones. Crucial in this respect is the degree to which fundamental material needs have been met during the formative period of individuals of a given generation. The post-materialism thesis, therefore, combines two hypotheses: a scarcity hypothesis and a socialization hypothesis. The scarcity hypothesis states that individual priorities reflect their socioeconomic environment: the greatest subjective value is placed on what is in relative short supply. The socialization hypothesis assumes that one's values reflect the conditions during one's pre-adult formative years and are relatively stable over the individual life course. Per Inglehart (1990), "Taken together, these two hypotheses imply that, as a result of the historically unprecedented prosperity and the absence of war that has prevailed in Western countries since 1945, younger birth cohorts place less emphasis on economic and physical security than do older groups, who have experienced a much greater degree of economic insecurity and that, conversely, the younger birth cohorts tend to give a higher priority to nonmaterial needs,

such as a sense of community and the quality of life" (p. 56). Inglehart's post-materialism thesis is based on the well-known need hierarchy of Maslow, suggesting that material needs must be fulfilled before immaterial or "higher" (growth) needs such as self-actualization become important. As a consequence, support for post-materialist values comes from post-war generations, the highly educated, and parts of the middle class. By contrast, pre-war generations, the less-educated, and members of the old middle class and the working class are more likely to support materialist values.

2.3. Rokeach and Braithwaite on the Political Importance of Values

Directly relevant to the domain of political psychology are the contributions of Rokeach and Braithwaite. Rokeach developed the two-value model of political ideology. He assumed that political ideologies differ from each other in terms of the social values freedom (independence, free choice) and equality (brotherhood, equal opportunity for all): the combination of these values enables us to describe and contrast all ideologies. According to Rokeach, socialism emphasizes both freedom and equality, whereas fascist ideology values neither freedom nor equality. Capitalism and communism both emphasize one of these core values more than the other: capitalism values freedom at the expense of equality, and communism values equality at the expense of freedom. Support for this model was found through a values content analysis of the writings of crucial authors advocating these different ideologies. Analyses of the values of supporters of different ideologies (such as voters or candidates of political parties), however, are less consistent with Rokeach's theory. This research generally shows rather important differences in the importance attributed to equality, with leftist activists showing a greater emphasis on equality than right-wing or conservative supporters. This is in line with Rokeach's theory. The value of freedom does not discriminate these different ideological groups, however, which refutes his claim.

These findings inspired Braithwaite to elaborate Rokeach's two-value model. She reduced a large number of values to two broader value orientations. The first is called "international harmony and equality," and covers values such as participation, cooperation, peace, and equal opportunities for all. The second refers to "national strength and order," and covers values such as national security, economic advancement, order, and security. Whereas the value orientation international harmony and equality captures Rokeach's notion of equality, national strength and order refers to a concept different than freedom. Braithwaite also shows that both value orientations are highly relevant for political attitudes and behaviors. Those favoring international harmony and equality emphasized progressive political views, expressed a preference for a leftist political party, and were more inclined toward political activity. The opposite holds for supporters of national strength and order: they preferred conservative views and a right-wing political party and were less inclined to engage in political activity.

2.4. Schwartz and Triandis: More Encompassing Value Orientations

Next to the previous contributions, two broad, more encompassing theories on value orientations have been developed, integrating and expanding previous research in the area. Building on the work of Rokeach, Schwartz proposed a universal theory of the content and structure of values. This theory has been tested in more than 60 countries, and is proven to be cross-culturally stable. Rokeach's value survey comprised 36 values, which were presented in two groups: 18 terminal values and 18 instrumental values. Respondents were asked to rank the values within each list according to their own preference. Schwartz increased this list to 56 values in his value inventory, and asked respondents to score each value on a 9-point scale. So, in addition to the increase in the number of values (again split up into two series), he also changed Rokeach's method of scoring from "ranking" to "rating," enabling the data to be analyzed with more sophisticated statistical methods. Schwartz divided the 56 values in his survey into ten motivational value types, according to the type of motivational goal they express. Table I summarizes these value types and their content.

An inspection of Table I shows that Schwartz's ten value types cover all the different value orientations discussed in the preceding section. Kohn's dichotomy between self-direction and conformity is covered by identical labels (with identical content), whereas Rokeach's core value freedom is part of self-direction. Universalism is closely related to Inglehart's post-materialism and Braithwaite's international harmony and equality, and includes Rokeach's core value of

TABLE I
Schwartz's Ten Motivational Value Types

Value type	*Definition*	*Exemplary values*
Power	Social status and prestige, control, or dominance over people and resources	Social power, authority, wealth
Achievement	Personal success through demonstrating competence according to social standards	Successful, capable, ambitious
Hedonism	Pleasure and sensuous gratification for oneself	Pleasure, enjoying life
Stimulation	Excitement, novelty, and challenge in life	Daring, varied life, exciting life
Self-direction	Independent thought and action—choosing, creating, exploring	Creativity, curious, freedom
Universalism	Understanding, appreciation, tolerance, and protection for the welfare of all people and for nature	Broad-minded, social justice, equality, protecting the environment
Benevolence	Preservation and enhancement of the welfare of people with whom one is in frequent personal contact	Helpful, honest, forgiving
Tradition	Respect, commitment, and acceptance of the customs and ideas that traditional culture or religion provide	Humble, devout, accepting my portion in life
Conformity	Restraint of actions, inclinations, and impulses likely to upset or harm others and violate social expectations or norms	Polite, obedient, honoring parents and elders
Security	Safety, harmony, and stability of society, of relationships, and of self	National security, social order, clean

Source: Schwartz (1994).

equality. Finally, security covers both Inglehart's materialism and Braithwaite's national strength and order. Schwartz, however, distinguishes more value types than these already suggested ones, thus enlarging the scope and possibilities of values research.

Schwartz represents his structure and value types by means of multidimensional scaling. This technique represents the values as points in a multidimensional space. The spatial distance between the points represents the relationship between the values. The more the ratings of a value correlate, the closer the values are located in the graphical representation. For Schwartz's universal values theory, this technique implies that the separate values of his motivational types must lie in each others' proximity and must be distinguishable as a separate field or region in the graphical representation. In contrast, values and value types that express opposing motivations should be discriminated clearly from one another. Fig. 1 shows the value structure obtained by Schwartz by means of such an analysis.

Schwartz claims that the oppositions in Fig. 1 can be classified into two bipolar dimensions. He describes the first dimension as openness to change, as opposed to conservation. This dimension contrasts value types such as stimulation, and self-direction (openness to change) with value types such as security, conformity, and tradition (conservation). The second dimension is formed by the opposition between self-transcendence and self-enhancement. Value types such as universalism and benevolence are opposed to value types that refer to power and achievement. Schwartz's subsequent research shows that his model and the two abstract dimensions are (very nearly) universal.

Schwartz's use of both higher order dimensions fits in with a research school that classifies the structure of values on the basis of two dimensions, giving rise to four value orientations. The key scholar in this school is Triandis. The first dimension distinguished by Triandis refers to the opposition between individualism (focus on the self) and collectivism (focus on others). The value types tradition, benevolence, and conformity correspond to collectivism. Values such as hedonism and stimulation refer to individualism. The opposition between "vertical" and "horizontal" thinking constitutes the second dimension. Horizontalism means that people consider each other as equal. Universalism typifies this orientation. Vertical thinking accepts forms of inequality, hierarchy, and authority (cf. the value type power). Triandis' two fundamental dimensions allow a more unambiguous interpretation

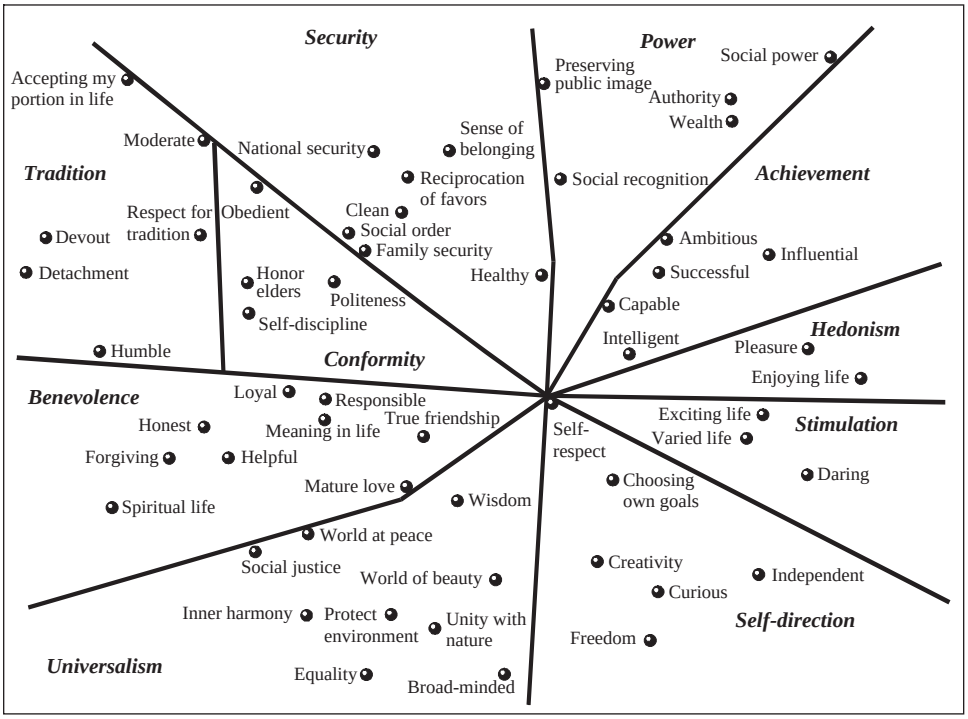


FIGURE 1 Visual representation of Schwartz's value structure. From Schwartz (1994).

than the two abstract value dimensions advanced by Schwartz. Combining the views of both scholars leads to four value orientations:

1. Horizontal collectivism or universalism combines the idea of equality and the focus on others to form a universalistic, altruistic value pattern. Equality, social justice, and true friendship are key values in this orientation. This value orientation appears in the bottom left in Fig. 1.

2. Horizontal individualism or self-direction combines the idea of equality with a focus on the individual. This leads to an emphasis on autonomy (with values such as choosing one's own goals and freedom), with peaks such as stimulation (daring and exciting life) and (slightly less central) hedonistic values (such as pleasure and enjoying life). This value orientation is located at the bottom right in Fig. 1.

3. Vertical individualism or competition combines the acceptance of inequality (the existence of a hierarchically ordered society) with a focus on the individual. The central idea is that the individual can achieve the "top" in society by adopting a competitive attitude. Central values are the commitment to achieving wealth, social recognition, and social power. This value orientation is located at the top right in Fig. 1.

4. Vertical collectivism or conformity, finally, combines a focus on others with the acceptance of a hierarchy. Supporters of this view are prepared to submit themselves to the collective. Key values are obedience, politeness and respect for tradition. It is located at the top left in Fig.1.

3. ON THE DIMENSIONALITY OF (CONSERVATIVE) IDEOLOGY

Values are the building stones of ideologies. The exact content of contemporary ideology, however, has been the subject of large-scale debates between scientists of different disciplines. Even though important differences in position exist, many scholars agree that (social and political) ideology can be captured under the heading of conservatism. The definition of conservatism, however, again proves difficult. Some, like Wilson (1979), proposed a one-dimensional concept, caused by "a generalized susceptibility to feeling threat or anxiety in the face of uncertainty" (p. 65). More prominent are multidimensional accounts. Van Deth and Scarbrough distinguished three important broad

value orientations (or ideologies), which cover important aspects of advanced industrial society as well as the legacies of industrial and pre-industrial societies. Inglehart's postmaterialism (as opposed to materialism) was selected as crucial for advanced (post-)industrial society. The distinction between left and right regarding economic issues was selected as a central value orientation reflecting the class conflicts typical of industrial society. Religious–secular value orientations were their third choice, referring to pre-industrial sources of political conflict.

3.1. Two Dimensions of Conservatism

Reviewing the field nevertheless makes clear that distinguishing a two-dimensional concept of ideology offers many advantages. This option is, first of all, suggested by several important scientists from different disciplines, such as Eysenck in psychology, and Lipset in political sciences. More importantly, these authors arrived at a rather similar structure, which corresponds closely to the thorough conceptualization of conservatism developed by the Dutch scholar Middendorp. On the basis of historical and theoretical considerations and a large body of empirical research, Middendorp distinguished two dimensions within conservatism and revealed their underlying values: freedom versus equality. The first was labeled socio-economic left-right and the second socio-cultural libertarianism-authoritarianism. In accordance with the literature in this field, these dimensions can be relabeled as economic conservatism (versus progressiveness) and sociocultural conservatism (versus progressiveness). Both dimensions are relatively stable over time, and—more surprisingly—are unrelated: respondents can combine a progressive view on one dimension with a conservative view on the other. Research also suggests that this two-dimensional conceptualization is rather universal.

The opposition between progressive and conservative attitudes in the socioeconomic domain points to socioeconomic cleavages in industrial society: conflicts between labor and capital. Middendorp showed empirically that the economic opposition between liberalism (in its European meaning of the concept) and socialism lies at the basis of this dimension: those who express progressive attitudes reject social inequality as unfair and unacceptable and argue for government intervention in the economy and for a more active trade union policy in order to reduce economic inequality. Conservatives express opposition to government intervention in the economy and against (militant) trade unions. They equally reject the claim that social differences are caused by inequality of opportunity, and emphasize private initiative and competition between individuals as essential for proper economic functioning.

The opposition between conservative and progressive on a sociocultural level relates to societal norms and traditions with respect to individual private lifestyle and family life. Ethical topics (such as sexuality and bioethics) and more social themes (such as free speech and tolerance toward minorities) belong to this field. This ideological cleavage relates to Van Deth and Scarbrough's religious versus secular value orientation, typical of pre-industrial sources of political conflict. Cultural conservatism thus can be considered as a relict of the old cleavage between the Church and the State. Historically, the non-religious community in particular has opposed the norms and values dictated by the Church, whereas the latter promoted traditional norms and conformity to traditional institutions. In a conservative sociocultural view, traditional norms in society are emphasized: adherents plead to maintain existing social norms and traditions ("traditionalism"). They prefer the traditional division of roles between men and women and support severe child-rearing practices in which strictness and obedience are stressed. They also defend the traditional work ethic, which emphasizes the moral duty to work. Those who take a progressive stand in the sociocultural field generally hold opposite attitudes. Traditional norms and attitudes are put into perspective. They stress the emancipation of women and the independence of children, support the liberalization of abortion and euthanasia, and emphasize other reasons for work than strict duty-related ones (e.g., self-actualization).

3.2. Relationship with Values and Value Orientations

These two broad ideological dimensions are of course related to values. Middendorp argued that Rokeach's values freedom and equality are at the core of both conservatism dimensions. On the socioeconomic level, progressive thought emphasizes equality: unequal opportunities and an unequal distribution of welfare and well-being are rejected. In conservative thought, the emphasis on the socioeconomic level is on the value of (individual) freedom: private initiative and competition among individuals should not be limited. On the sociocultural level, this emphasis is

reversed. The progressive way of thinking emphasizes the freedom of the individual to arrange life according to one's own insight (self-direction). In the conservative way of thinking, such individual freedom is rejected and equality is stressed. Conservatives on the sociocultural level emphasize that everyone should adhere to the traditional norms and values in society (conformity and traditionalism). So, individuals who emphasize the basic value of individual freedom adhere to opposing views on both ideological dimensions. On the economic level, they show conservative preferences, whereas on the sociocultural level they choose progressive viewpoints. The reverse is true for those emphasizing equality. They combine a progressive view on economic issues with a conservative one regarding sociocultural issues.

Both ideological conservatism dimensions are related not only to separate values, but also to larger value dimensions. Sociocultural conservatism is related to Kohn's distinction between self-direction and conformity. Those who emphasize conformity to external authorities find it equally important to obey the norms and traditions in society (traditionalism). These individuals stress external standards for their behavior: "others" have to tell us what to do. These others can be superiors in the hierarchy (vertical conformity), as well as equals (horizontal conformity). Individuals who take a progressive position on sociocultural issues defend freedom of thought and conscience. They emphasize the right to organize their lives according to their own insights, even if it means deviation from current norms. This view is closely related to self-direction, since the norm for one's own behavior is sought with oneself (internal standards for behavior).

Economic attitudes are related to value orientations as well. The value opposition between universalism and competition correlates with economic conservatism: economic progressiveness shows a positive correlation with the value orientation universalism and a negative correlation with the value orientation competition. The reverse is true for economic conservatism (positive correlation with competition and negative correlation with universalism).

3.3. View on Human Nature and Society, and Relationship with Social Groups

Earlier, ideology was defined as a (relatively) cohesive world view and portrayal of humans, reflecting the interests of a specific social group. How does our treatment of conservatism relate to these aspects? Based on an historical analysis, Middendorp showed that the ideology of conservatism combines a rather negative view on human nature with a rather positive view on the nature of society. In conservatism, one emphasizes that people are inclined to egoism rather than to altruism. As a consequence, human beings have to be restricted and disciplined by society. Societal institutions, on the other hand, incorporate valuable traditions and wisdom, which enable human beings to behave in a civilized way. This logic, however, is an elite construction. Much knowledge of political and historical trends and a high level of conceptual and cognitive sophistication are needed to understand its logic. As a consequence, research shows that only members of the political elite (e.g., members of parliament) are capable of integrating both conservatism dimensions into one single dimension, ranging from conservatism to progressiveness.

This is not the case within the normal population. "Ordinary" citizens distinguish the two forms of conservatism discussed previously. Both forms of conservatism are related to specific groups, reflecting their interests. Low-income groups are prompted toward an economic progressive position, since this view promotes their interests. Groups with a higher level of education are prompted to stress progressive views regarding sociocultural issues, since these views enable them to organize their lives according to their own insights. Because cultural conservatism is linked to the cleavage between the Church and the State, religious respondents tend to stress this form of conservatism to a larger degree than do non-religious people.

4. POLITICAL CONSEQUENCES OF IDEOLOGY AND VALUE ORIENTATIONS

Ideology and values have important consequences for political behaviors of citizens. Before one can act at the political level, however, one must feel motivated to do so. Political interest and political efficacy play an important role in translating ideology and values into political action. Comparative research shows postmaterialism and (to a lesser extent) cultural progressiveness to be associated with political interest. A similar finding applies to political efficacy: the self-image that one is an effective and influential participant in politics. Efficacious citizens see themselves

as comprehending, controlling, and mastering their political environments. People holding post-materialist values perceive themselves to be more efficacious and subjectively competent. This of course strengthens their wish to participate in the political process. Also, cultural progressiveness enhances people's sense of civic competence and political efficacy. However, the net impact of value orientations on political interest and political efficacy is modest, and other variables, such as the level of education, play a more direct and important role.

Given the results mentioned above, it comes as no surprise that post-materialism (and, to a lesser extent, cultural progressiveness) is found to influence different forms of unconventional protest activity, such as signing petitions, attending lawful demonstrations, and participating in boycotts or occupations of buildings. The motivating force of ideology is not limited to these unconventional forms of activity. Research shows that trade union participation is also linked to ideology and values. Economic progressiveness in particular seems relevant in this context, because it is associated with participation that requires a high level of effort, such as becoming a union activist or participation in industrial action (such as strikes). However, even though informal political activities became important in many countries, voting is still the most frequent mode of political participation. Interestingly, research shows that post-materialist values have little impact on electoral turnout levels.

The impact of ideology and values is much more substantial when one analyzes party choice. Here, ideological orientations play a rather important role in distinguishing the electorates of different parties. In analyzing the Dutch vote, Middendorp plotted all parties according to the mean score of cultural and economic conservatism of their voters. This results in an ordering of all parties on the economic dimension (the "left-right" vote) and on the cultural one (the "libertarian-authoritarian" vote). He concluded that the left-right vote is mostly explained by economic conservatism, social class, and income. The libertarian-authoritarian vote was mainly explained by cultural conservatism and religious involvement. A more encompassing analysis of different countries shows that these findings can be generalized across Europe. Socialist and communist parties (e.g., "leftist" parties) have the most economic progressive electorates. The electorates of the (economic conservative, "rightist") liberal parties are the most rightist in each country in continental Europe. Christian democratic voters tend to be centrists, leaning toward the center right of economic conservatism. Green parties and left-wing socialist parties have the most post-materialist electorates in most countries. Conservative and Christian democratic electorates in most countries are predominantly materialist. In this comparative research, cultural conservatism was not included, so no conclusions regarding this ideology could be reached. Other research, however, suggests that this ideological dimension too is relevant in determining the political vote.

Interestingly, different scholars find that ideological dimensions explain more variance than background characteristics do. The discriminating power of variables such as social class and religion tends to decline somewhat over time, whereas the impact of values on voting intentions seems rather stable in most countries or increases slightly when compared to a more distant past. Some consider this finding an indication of the "ideologization" of the vote: ideology seems to determine the vote to a larger extent in recent elections than in prior ones. Other scholars suggest that instead we are witnessing the rise of issue voting. According to issue salience theory, voting is determined by those value orientations and issues that dominate the political agenda and party conflicts in given political systems and at a specific time. Because these issues are variable, the vote becomes more unpredictable than before, which explains the higher proportion of volatile voters, changing their party preference in between elections. Issues, however, are not unrelated to ideology or value orientations, but are embedded in more encompassing ideological orientations. This means that ideology is still relevant in issue salience theory, even though particular political circumstances might single out specific aspects of this ideology. A good example of this trend is the rise of extreme-right wing parties in Europe. Research consistently shows that these parties appeal to the negative attitude of citizens towards foreigners. Extreme right-wing parties have been successful in putting their anti-immigrant stands on the political agenda, and successfully persuaded the electorate that they will deal with this topic once elected. Negative attitudes toward foreigners are not an isolated domain, however, but are firmly rooted in cultural conservatism. The rise of extreme-right wing parties in Europe thus suggests that at least a part of this kind of conservatism grew in importance as predictor of voting behavior, compared to the past.

See Also the Following Articles

Political Psychology, Overview ■ Values and Culture

Further Reading

Borg, S. (1995). Electoral participation. In J. W. Van Deth, & E. Scarbrough (Eds.), *The impact of values* (pp. 441–460). Oxford: Oxford University Press.

Braithwaite, V. (1994). Beyond Rokeach's equality-freedom model: Two-dimensional values in a one-dimensional world. *Journal of Social Issues*, *50*, 67–94.

De Witte, H. (1996). Are trade union members (still) motivated by ideology? A review of the importance of ideological factors as determinants of trade union participation in (the Flemish part of) Belgium. In P. Pasture, J. Verberckmoes, & H. De Witte (Eds.), *The lost perspective? Trade unions between ideology and social action in the New Europe. Volume 2. Significance of ideology in European trade unionism* (pp. 275–304). Aldershot: Avebury.

De Witte, H., & Scheepers, P. (Eds.). (1999). *Ideology in the low countries. Trends, models and lacunae.* Assen: Van Gorcum.

Evans, G., Heath, A., & Lalljee, M. (1996). Measuring left-right and libertarianism-authoritarian values in the British electorate. *British Journal of Sociology*, *47*, 93–112.

Eysenck, H. J. (1972). *Psychology is about people*. London: Penguin Press.

Gabriel, O. W. (1995). Political efficacy and trust. In J. W. Van Deth, & E. Scarbrough (Eds.), *The impact of values* (pp. 357–389). Oxford: Oxford University Press.

Gabriel, O. W., & Van Deth, J. W. (1995). Political interest. In J. W. Van Deth, & E. Scarbrough (Eds.), *The impact of values* (pp. 390–411). Oxford: Oxford University Press.

Grunberg, G., & Schweisguth, E. (1993). Social libertarianism and economic liberalism. In D. Boy, & N. Mayer (Eds.), *The French voter decides* (pp. 45–64). Ann Arbor, MI: The University of Michigan Press.

Gundelach, P. (1995). Grass-roots activity. In J. W. Van Deth, & E. Scarbrough (Eds.), *The impact of values* (pp. 412–440). Oxford: Oxford University Press.

Inglehart, R. (1977). *The silent revolution. Changing values and political styles among Western publics*. Princeton: Princeton University Press.

Inglehart, R. (1990). *Culture shift in advanced industrial society*. Princeton, NJ: Princeton University Press.

Inglehart, R. (1997). *Modernization and postmodernization*. Princeton, NJ: Princeton University Press.

Johnson, S., & Tamney, J. (2001). Social traditionalism and economic conservatism: two conservative political ideologies in the United States. *The Journal of Social Psychology*, *141*, 233–243.

Knutsen, O. (1995). Party choice. In J. W. Van Deth, & E. Scarbrough (Eds.), *The impact of values* (pp. 461–491). Oxford: Oxford University Press.

Kohn, M. L. (1977). *Class and conformity: a study in values. With a Reassessment (2nd ed.). Chicago, IL: The University of Chicago Press.*

Kohn, M. L. (1995). Social structure and personality through time and space. In P. Moen, G. Elder, *et al*. (Eds.), *Examining lives in context: Perspectives on the ecology of human development* (pp. 141–168). Washington, DC: American Psychological Association.

Kohn, M. L., & Schooler, C. (1983). *Work and personality: An inquiry into the impact of social stratification*. Norwood, NJ: Ablex Publishing Corporation.

Lipset, S. M. (1981). *Political man, the social basis of politics*. Baltimore, MA: Johns Hopkins University Press.

Lubbers, M., Gijsberts, M., & Scheepers, P. (2002). Extreme right-wing voting in Western Europe. *European Journal of Political Research*, *41*, 345–378.

Middendorp, C. (1978). *Progressiveness and conservatism. The fundamental dimensions of ideological controversy and their relationship to social class*. The Hague/Paris/New York: Mouton.

Middendorp, C. (1991). *Ideology in Dutch politics: The democratic system reconsidered 1970–1985*. Assen: Van Gorcum.

Rokeach, M. (1973). *The nature of human values*. New York: The Free Press.

Rokeach, M. (1979). *Understanding human values: Individual and societal*. London: Collier MacMillan.

Scheepers, P., Lammers, J., & Peters, J. (1994). Religious and class voting in The Netherlands 1990–1991: A review of recent contributions tested. *The Netherlands' Journal of Social Sciences*, *30*, 5–24.

Schwartz, S. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In M. Zanna (Ed.), *Advances in experimental social psychology*, Vol. 25 (pp. 1–65). New York: Academic Press.

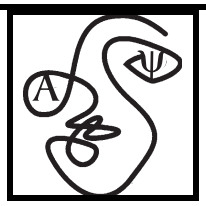
Schwartz, S. (1994). Are there universal aspects in the structure and content of human values? *Journal of Social Issues*, *50*, 199–45.

Triandis, H. (1995). *Individualism and collectivism*. Boulder, CO: Westview Press.

Van Deth, J. W., & Scarbrough, E. (Eds.). (1995). *The impact of values*. Oxford: Oxford University Press.

Van Gyes, G., & De Witte, H. (1999). Conservatism and values. A confrontation of Middendorp with Schwartz and Triandis. In H. De Witte, & P. Scheepers (Eds.), *Ideology in the low countries. Trends, models and lacunae* (pp. 121–150). Assen: Van Gorcum.

Wilson, G. (1979). *Manual for the Wilson-Patterson Attitude Inventory*. Windsor: NFER.



Indecision, Vocational

Paul J. Hartung
Northeastern Ohio Universities College of Medicine, Rootstown, Ohio, USA

1. Perspectives on Indecision
2. A Brief Historical Synopsis: From Plato to Parsons to the Present
3. Vocational Indecision: The Positivist Perspective
4. Vocational Indecision: The Phenomenological Perspective
5. Future Directions
Further Reading

GLOSSARY

indecisiveness A global problem in making life decisions, trait anxiety associated with decision making, and chronic difficulty in resolving the decisional process.

logical positivism A philosophy of science that advances use of the scientific method to produce knowledge about what are believed to be discoverable facts in the world; it emphasizes objectivity and empirical data.

phenomenology A philosophy of science that views knowledge and meaning as derived from within the individual interpreter of experiences, events, and observations; it emphasizes subjectivity and intuitive meaning.

vocational indecision The inability to express an educational or occupational choice when asked to do so; it involves a delay in bringing closure to the career choice process, that is, a wavering, pause, or hesitation in vocational development and an openness to alternative career pathways.

Vocational indecision remains a core and complex construct in contemporary career psychology, linked intimately to the career choice and decision-making process. It also persists among the most common concerns presented in career counseling and has served as the focus of a vast literature of empirical research and conceptual analysis. Historically, vocational indecision has denoted an inability to express an educational or occupational choice when asked to do so and a delay in bringing closure to the career choice process. Contemporary definitions further characterize indecision as a wavering, pause, or hesitation in vocational development and an openness to alternative career pathways. A synthesis of traditional and alternative interpretations enriches the meaning and the management of vocational indecision.

1. PERSPECTIVES ON INDECISION

A scheme useful for both elaborating on and dealing with the complexity of vocational indecision is to consolidate two principal perspectives on the construct that have emerged during the course of comprehending and intervening to address it. One perspective perceives indecision as an objective circumstance, whereas the other perspective views indecision as a subjective experience. An encapsulated history of career decision making, and indecision as a component of the decision-making process, provides a broader context for discussing these two perspectives in turn.

©2004 Elsevier Inc.
All rights reserved.

2. A BRIEF HISTORICAL SYNOPSIS: FROM PLATO TO PARSONS TO THE PRESENT

Inquiry into career decision making in the Western world dates back to the Greek philosopher Plato, who assisted individuals with their vocational decisions by proffering wisdom, advice, and recommendations about fitting work. In the East, teachings in the Tao Te Ching have long advised individuals to do what they enjoy in work. Before the dawn of the early 20th-century vocational guidance movement, other people (e.g., Lysander Richards) used quasi-scientific methods of graphology, physiognomy, palmistry, and phrenology to identify suitable work for people.

With the establishment in 1909 of the Vocation Bureau in Boston by Frank Parsons and his colleagues came an approach to understanding and alleviating vocational indecision grounded in logical analysis and scientific inquiry. By advocating the use of the scientific method, Parsons championed the positivist perspective on vocational indecision that predominated throughout the 1900s and continues to hold great sway in vocational psychology and career counseling today. An alternative perspective on vocational indecision, initiated by D. E. Super and his thematic extrapolation method at midcentury, advanced by D. V. Tiedeman and his colleagues during the 1960s, and propelled by prevailing constructivist and narrative approaches to career decision making, uses phenomenology to comprehend the individual experiences of decision makers. The positivist and phenomenological perspectives both contribute uniquely to comprehending and coping with vocational indecision.

3. VOCATIONAL INDECISION: THE POSITIVIST PERSPECTIVE

The traditional perspective on vocational indecision derives from logical positivism and views indecision as an objective phenomenon. It prizes the use of psychometric scales to yield scores that indicate degree and type of vocational indecision. Three distinct phases of this perspective spanning the past century have transformed indecision from a simple dichotomous concept into a complex multidimensional one.

3.1. A Dichotomy

Parsons initially defined indecision for vocational psychology as a dichotomy consisting of two groups. One group contained individuals decided about their educational or vocational plans, whereas the other group contained persons undecided about their educational or vocational plans. This rudimentary scheme (a) framed indecision as a problem, (b) sparked 40 years of empirical inquiry into personality and situational differences between members of these two groups, (c) characterized undecided persons as deficient if not pathological, and (d) engendered counseling efforts to treat the personality defect underlying the indecision. With the advent of the developmental perspective on careers during the 1950s, some theorists began to advance a view of indecision as a normal component of human development from which a career-decided state would emerge. This view preceded contemporary perspectives on indecision as a healthy state of adaptive uncertainty, openness to alternative career possibilities, and flexibility in coping with change and chance events.

3.2. A Unidimensional Continuum

Research based on a dichotomous concept of indecision produced mixed and inconsistent findings, with many studies supporting differences between decided and undecided individuals and many other studies indicating no differences between these groups. Two approaches, one psychometric and the other theoretical, were taken to address these inconsistencies. The psychometric approach sought to better measure career indecision and career decision-making status. The theoretical approach endeavored to better conceptualize what career indecision means.

Taking a psychometric approach, researchers during the 1970s turned to a view of indecision as a unidimensional continuum ranging from undecided to decided. Psychometric scales, such as the Vocational Decision-Making Difficulties Scale and the Career Decision Scale, were subsequently constructed to index positions on this continuum. So conceived, indecision reflected a continuous process operationally defined by scale scores indicating level of decidedness and difficulties that thwart choice.

3.3. A Multidimensional Construct

Advancing the psychometric approach to indecision, the objective perspective has evolved to the current

view of indecision as a multidimensional concept with distinct and measurable subtypes. Rather than perceiving all undecided persons as undecided for the same reasons, the multidimensional perspective recognizes heterogeneous undecided subgroups. Researchers consequently turned to either revising existing scales or developing new measures to assess degree of and reasons for indecision, which would then guide intervention. These measures now constitute two generations of instrument development. First-generation measures operate on a unidimensional definition of indecision and view groups of individuals with equivalent overall scale scores as homogenous. Second-generation measures construe indecision as multidimensional and seek to identify subgroups of undecided individuals and provide different interventions for each type.

3.4. Indecision Versus Indecisiveness

During the 1960s, a theoretical approach to resolving inconsistent research findings about differences between decided and undecided groups centered on distinguishing indecision from indecisiveness. Leona Tyler proposed that indecisiveness reflects a global problem in making life decisions, trait anxiety associated with decision making, and chronic difficulty in resolving the decisional process. Indecision, in contrast, reflects a specific problem to make an occupational choice, a state of situational anxiety about decision making, and a developmental delay in resolving the decisional process. This view on indecision has been propelled by subsequent theorizing about indecision as a natural component of career development, as an opportunity to explore, and as open-mindedness to random events and opportunities in shaping career-related decisions. The intuitively reasonable distinction between indecision and indecisiveness as differentially related to anxiety has met with little empirical support.

4. VOCATIONAL INDECISION: THE PHENOMENOLOGICAL PERSPECTIVE

Phenomenology offers an alternative to the positivist perspective on vocational indecision. Situated in constructivist and narrative approaches to career that have emerged more prominently during the past decade or so, the phenomenological view conceives of indecision as a subjective experience. Consequently, rather than using psychometric inventories and objective scores to measure level and type of indecision, the subjective perspective prizes the use of narratives and stories to reveal life themes, patterns, projects, and goals that guide movement within the choice process. From the phenomenological perspective, indecision reflects a wavering or pause in vocational development and an openness to alternative career pathways. Individuals can use indecision to contemplate their directions in life, construct possible futures, and convert their identities into truer self-representations.

The subjective perspective on indecision emphasizes the personal meaning ascribed to the lived experience of being undecided. Exploring undecided individuals' subjective, qualitative, and personal realities leads to comprehending their unique experiences of indecision, preoccupations and intentions, and possibilities for constructing their future lives. By eliciting narratives, reviewing early recollections and family stories, and situating work in the context of people's lives, phenomenologically based approaches involve a collaborative endeavor to imbue indecision with personal meaning. Vocational decision making becomes a process of untangling a present filled with uncertainty and doubt to clarify and envision a future path that will provide purpose, direction, fulfillment, and satisfaction.

5. FUTURE DIRECTIONS

Vocational indecision has long been a focus of theory, research, and practice in vocational psychology. Decades of scholarship have examined indecision both as an objective circumstance, with distinct measurable subtypes, and as a subjective experience, wholly unique to the individual who experiences it. Future efforts to meld positivist and phenomenological perspectives on indecision will likely yield a more comprehensive approach to comprehending the career decision-making process and to dealing with vocational indecision as a problem that can interrupt that process. Emerging perspectives on vocational indecision that argue against decidedness as the optimal vocational outcome in favor of blending planfulness, flexibility, and more active engagement with random opportunities for shaping future decisions will also expand possibilities and potential for human development through work and career.

See Also the Following Articles

Decision Making ■ Holland's Theory (Vocational Personality Types)

Further Reading

Cochran, L. (1997). *Career counseling: A narrative approach*. Thousand Oaks, CA: Sage.

Gati, I., & Asher, I. (2001). The PIC model for career decision making: Prescreening, in-depth exploration, and choice. In F. T. L. Leong, & A. Barak (Eds.), *Contemporary models in vocational psychology: A volume in honor of Samuel H. Osipow* (pp. 7–54). Mahwah, NJ: Lawrence Erlbaum.

Gelatt, H. (1989). Positive uncertainty: A new decision-making framework for counseling. *Journal of Counseling Psychology*, *36*, 252–256.

Grites, T. J. (1981). Being "undecided" might be the best decision they could make. *The School Counselor*, *29*, 41–46.

Hartung, P. J., & Blustein, D. L. (2002). Reason, intuition, and social justice: Elaborating on Parsons's career decision-making model. *Journal of Counseling and Development*, *80*, 41–47.

Krieshok, T. S. (1998). An anti-introspectivist view of career decision making. *Career Development Quarterly*, *46*, 210–229.

Krieshok, T. S. (2001). How the decision-making literature might inform career center practice. *Journal of Career Development*, *27*, 207–216.

Krumboltz, J. D. (1992). The wisdom of indecision. *Journal of Vocational Behavior*, *41*, 239–244.

Krumboltz, J. D. (1998). Serendipity is not serendipitous. *Journal of Counseling Psychology*, *45*, 390–392.

Mitchell, K. E., Levin, A. S., & Krumboltz, J. D. (1999). Planned happenstance: Constructing unexpected career opportunities. *Journal of Counseling and Development*, *77*, 115–124.

Phillips, S. D. (1997). Toward an expanded definition of adaptive decision making. *Career Development Quarterly*, *45*, 275–287.

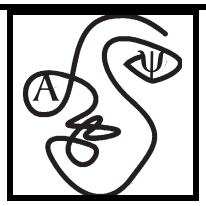
Phillips, S. D., & Pazienza, N. J. (1988). History and theory of the assessment of career development and decision making. In W. B. Walsh, & S. H. Osipow (Eds.), *Career decision making* (pp. 1–31). Hillsdale, NJ: Lawrence Erlbaum.

Savickas, M. L. (1995). Constructivist counseling for career indecision. *Career Development Quarterly*, *4*, 363–373.

Savickas, M. L. (2000). Assessing career decision making. In C. E. Watkins, & V. Campbell (Eds.), *Testing and assessment in counseling practice* (2nd ed., pp. 429–477). Mahwah, NJ: Lawrence Erlbaum.

Slaney, R. B. (1988). The assessment of career decision making. In W. B. Walsh, & S. H. Osipow (Eds.), *Career decision making* (pp. 33–76). Hillsdale, NJ: Lawrence Erlbaum.

Tinsley, H. E. (Ed.). (1992). Career decision making and career indecision [special issue]. *Journal of Vocational Behavior*, *41*(3).



Indigenous Psychologies

Uichol Kim *Chung-Ang University, Seoul, Korea*

Young-Shin Park *Inha University, Incheon, Korea*

1. Historical Background
2. What Is Indigenous Psychology?
3. Two Approaches in Indigenous Psychology
4. How Is Research Done?
5. An Example: Educational Attainment
6. Future Directions
   Further Reading

GLOSSARY

amae A Japanese term referring to the act of asking for and receiving special favors in close relationships.

chong A Korean term meaning deep affection and attachment for a person, place, or thing.

extrinsic motivation A factor outside an individual (e.g., parental influence, social support, social reward) that can motivate a person into action.

guanxi A Chinese term meaning relationship.

indigenization from within Developing theories, concepts, and methods internally, with indigenous information being considered as a primary source of knowledge.

indigenization from without Taking existing psychological theories, concepts, and methods and modifying them to fit the local cultural context.

indigenous psychology According to Kim and Berry, "the scientific study of human behavior [or mind] that is native, that is not transported from other regions, and that is designed for its people."

intrinsic motivation A factor within an individual (e.g., personal interest, preference, ability) that can motivate a person into action.

phenomenological knowledge Internal thoughts, feelings, and motives of a person.

practical validity Theories that can help one to understand, predict, and manage the world.

transactional model A scientific model in which phenomenological knowledge is systematically linked with the situation or event, on the one hand, and with behavior or performance, on the other.

Indigenous psychology represents the cultural psychology tradition espoused by Wundt that examines human action by linking people's intentions, skills, and goals with the situation or event, on the one hand, and with behavior or performance, on the other. It advocates studying people in their natural contexts, using concepts that are meaningful, and investigating the skills that people develop to manage their lives. The goal of indigenous psychology is to create more rigorous scientific theories that can be verified and applied in the natural setting. The phenomenal educational achievement of East Asian students is provided as an example to highlight this approach.

1. HISTORICAL BACKGROUND

Wundt is considered to be the "father of modern psychology." In establishing psychology as an independent scientific discipline, he recognized two traditions in science: the natural sciences tradition and the

©2004 Elsevier Inc.
All rights reserved.

cultural sciences tradition. Wundt was instrumental in establishing the experimental method to study basic physiological, perceptual, and cognitive processes. However, he recognized the limitations of experimental method and emphasized the need to develop cultural psychology. He observed that thinking is heavily conditioned by language and culture. He regarded cultural psychology as a more important aspect of psychological research. General psychology represents the natural sciences approach, and indigenous psychology represents the cultural science tradition.

When psychology became established in North America, psychologists adopted the natural science approach, and eliminated the cultural approach, in search of abstract universal laws of human behavior. Behaviorism emerged as the dominant paradigm, and the subject matter of psychology became tailored to fit into the narrow definition of natural science. Early psychologists attempted to eliminate subjective aspects (e.g., consciousness, intentions, meaning) and to conduct controlled experiments in laboratory settings. Psychologists assumed that the results of laboratory studies could be generalized to explain complex human behavior in the natural setting.

Indigenous psychologies evolved around the world as a reaction against unjustified claims of universality of psychological theory. Although existing psychological theories and concepts are assumed to be objective, value free, and universal, in reality they are deeply enmeshed with Euro-American values that champion rational, liberal, individualistic ideals. Indigenous psychology represents an alternative scientific paradigm in which the human qualities (e.g., meaning, goals, creativity) and culture have been integrated as central elements of research design.

2. WHAT IS INDIGENOUS PSYCHOLOGY?

In 1993, Kim and Berry defined indigenous psychology as "the scientific study of human behavior [or mind] that is native, that is not transported from other regions, and that is designed for its people." Indigenous psychology advocates examining knowledge, skills, and beliefs that people have about themselves and studying these aspects in people's natural contexts. It represents a bottom-up approach in which the goal of psychology is to understand how people function in their natural contexts and to examine how people interact with their natural and human worlds. Indigenous psychology recognizes that human psychology is complex, dynamic, and generative, and it recognizes the importance of agency, meaning, intentions, goals, and context in understanding behavior. Theories, concepts, and methods are developed to correspond with psychological phenomena. The goal is to create a more rigorous, systematic, universal science that can be verified both theoretically and empirically. In addition, there are seven facets of indigenous psychology that can be identified.

First, indigenous psychology emphasizes contextualized understanding rooted in a particular setting (e.g., ecological, political, historical, cultural). It emphasizes the discovery and use of natural taxonomies in search of regularities, general principles, and universal laws. It examines how people view themselves, relate to others, and manage their environment.

Second, contrary to popular misconception, indigenous psychologies are not studies of native peoples, ethnic groups, or people living in distant lands. Indigenous research has often been equated with anthropological analysis of "exotic" people living in distant lands. Indigenous psychology is needed for all cultural, native, and ethnic groups, including economically developed countries such as Canada, France, and the United States.

Third, acceptance of indigenous psychology does not affirm or preclude the use of a particular method. Indigenous psychology is part of the scientific tradition where an important aspect of the scientific endeavor is the discovery of appropriate methods for the phenomenon under investigation. Scientists should not, and cannot, be bound to a particular method. The use of multiple methods is recommended to increase one's confidence that a particular finding is valid and not an artifact of the research method. Results from multiple methods should be integrated to provide a more comprehensive and robust understanding of psychological phenomena.

Fourth, it has been assumed that insiders have a better understanding of indigenous phenomena and that outsiders can have only a limited understanding. Although a person who has been born and raised in a particular community may have insights and understanding of indigenous phenomena, this might not be true in all instances. As Wirth observed in 1946, "The most important thing . . . that we can know about a person is what he [or she] takes for granted, and the most elemental and important facts about a society are those things that are seldom debated and generally regarded as settled." An outsider, with an external point of view, can call attention to the fact that what is assumed to be natural is

actually cultural. Although an outsider may have a superficial understanding of indigenous phenomena found in other cultures, he or she may point out peculiarities, inconsistencies, and "blind spots" that insiders may have overlooked. Both internal and external points of view are necessary in providing a comprehensive and integrated understanding of psychological phenomena.

Fifth, many indigenous psychologists search philosophical and religious texts for explanations of indigenous phenomena. Too often, they use philosophical treatises (e.g., Confucian Classics) or religious texts (e.g., Koran, Vedas) as explanations of psychological phenomena. However, these texts were developed for philosophical and religious purposes several thousand years ago and not for psychological research in the 21st century. To use these texts, one must first translate them into psychological concepts and empirically verify how they influence the ways in which people think, feel, and behave. One cannot assume that because a person is Chinese, he or she will automatically follow the Confucian way, nor can one assume that Hindu Dharma can explain the behavior of an individual just because he or she is an Indian.

Sixth, as with other scientific traditions, one of the goals of indigenous psychology is the discovery of universal facts, principles, and laws that could explain human diversity. However, psychological universals must be theoretically and empirically verified rather than assumed *a priori*.

Seventh, human consciousness, agency, meaning, and goals are considered to be central constructs in explaining human behavior (or mind). The separation of the subjective and objective worlds that is important in the natural sciences is not possible in the human world because people are both the subject and the object of investigation. Researchers must recognize that research is a value-laden enterprise in search of a probabilistic understanding of human action rather than a search for objective and mechanical knowledge. Research topics and stimuli must be meaningful and contextualized rather than arbitrarily selected, manipulated, and controlled.

3. TWO APPROACHES IN INDIGENOUS PSYCHOLOGY

In 1993, Enriquez identified two approaches in indigenous psychology: indigenization from without and indigenization from within. Indigenization from without involves taking existing psychological theories, concepts, and methods and modifying them to fit the local cultural context. The approaches advocated by cultural and cross-cultural psychologists are examples of indigenization from without. In this approach, rather than assuming that a particular theory is universal *a priori*, researchers modify and adapt psychological theories and integrate them with the local cultural knowledge. Those aspects that can be verified across cultures are retained as possible cultural universals. Existing theories in cognitive, developmental, social, and organizational psychology have been modified and extended by indigenous research.

In the indigenization from within approach, theories, concepts, and methods are developed internally and indigenous information is considered to be a primary source of knowledge. For example, one of the core values and assumptions that psychologists from East Asia question is the emphasis on individualism. In East Asia, the relatedness of one person to another person is considered to be fundamental. In East Asia, the word for human being can be literally translated as "human between." In other words, it is what happens between individuals that makes them human. In Chinese culture, the concept of *guanxi* (literally "relationship") occupies center stage, defining interactions in family, school, companies, and society. In Japan, the concept of *amae* (literally "the act of asking and receiving special favors in close relationships") helps to define and maintain close interpersonal relationships. In Korea, the concept of *chong* (literally "deep affection and attachment for a person, place, or thing") is a central emotion that binds family, friends, and colleagues together. Recently, both cross-sectional and longitudinal studies yielded results that confirm the importance of close relationships that are highly reliable, valid, and applicable in Korea.

4. HOW IS RESEARCH DONE?

Indigenous psychology advocates development and use of methods that are appropriate for psychological phenomena. It recognizes the existence of two types of knowledge: objective third-person knowledge (i.e., analytical, semantic, and declarative knowledge) and subjective first-person knowledge (i.e., phenomenological, episodic, and procedural knowledge). Analytical, semantic, and declarative knowledge represents information based on objective and impartial analysis. Phenomenological, episodic, and procedural knowledge represents experiential knowledge that people have about themselves and ways in which to deal with their physical and human worlds.

In natural sciences, the goal is to obtain objective knowledge through rigorous experimentation, observation, and mathematical analysis. In natural sciences, phenomenological knowledge (e.g., the thoughts, feelings, and motives of a tree, lion, or star) represents irrelevant research questions. For example, if a lion kills a deer, one assumes that the behavior is caused by, and explainable in terms of, instinct. However, if a woman kills a man, one does not simply attribute the killing to her instinct. One seeks additional information such as motive, agency, and responsibility. The same act (i.e., a woman killing a man) has different meanings and consequences based on one's knowledge of the person's motive and agency. If a woman killed her husband to collect an inheritance, it is premeditated murder. If a woman found her husband in bed with another woman and killed him out of rage, it can be classified as manslaughter. If a woman killed a man while she was driving drunk, it can be classified as involuntary manslaughter. If a woman killed a man while resisting being raped, it can be classified as self-defense. If a 5-year-old girl kills her friend while playing with her father's gun, it can be classified as gross negligence on the part of the father. Based on one's assessment of the person's intentions and responsibility, one assigns different outcomes, ranging from capital punishment (in the case of premeditated murder), to 10 years imprisonment (in the case of manslaughter), to no punishment (in the case of self-defense).

In human sciences, unlike in physical and biological sciences, it is essential to know the intentions, goals, and motives of people. This information can be obtained through self-reports. In the transactional model espoused by both Bandura and Kim, the subjective human qualities (e.g., agency, intentions, meaning, goals) can be systematically linked with the situation or event, on the one hand, and with behavior or performance, on the other.

In research, phenomenological, episodic, and procedural knowledge needs to be translated into analytical knowledge. However, many people do not possess the analytical skills in describing how they performed certain actions. For example, adult native English speakers can freely express their thoughts in English (i.e., procedural knowledge), but they might not know the grammatical syntax or structure of the spoken words (i.e., semantic knowledge). In other words, they know how to produce the sentences, but they might lack the ability to describe them analytically. This is the case because people are taught to use the language to communicate with others. Grammatical knowledge is taught as a part of formal education in school. Conversely, Koreans might know the grammar of English, but they might have difficulty in speaking English because it is a different type of knowledge. It is the role of the researcher to analyze the grammar of psychological phenomena and to verify it to ensure that the analysis is correct. In human life, both experiential knowledge (e.g., a football player describing his experiences in playing a game) and analytical knowledge (e.g., a sport commentator providing a play-by-play analysis) are useful types of information that need to be integrated (e.g., a coach planning strategy for the next game). Indigenous psychological analysis has been conducted in the areas of achievement, conception of self, parent–child relationships, school violence, stress, occupational safety, trust, organizational behavior, and political culture. An example of educational achievement of East Asian students, focusing on Korean results, is provided in the next section.

5. AN EXAMPLE: EDUCATIONAL ATTAINMENT

In international comparisons of academic literacy of high school students (e.g., Third International Mathematics and Science Study [TIMSS], Program for International Student Assessment [PISA]), East Asian students are the top achievers in mathematics and science. Although the U.S. government spends more money per student to educate its children and adolescents, American schools have smaller classes, and many psychologists provide their expertise in developing U.S. educational programs, American students perform far below their East Asian counterparts. In the TIMSS, American students were ranked 19th in mathematics and 18th in science; in the PISA, they were ranked 15th in reading literacy, 19th in mathematical literacy, and 14th in scientific literacy. In contrast, East Asian students were ranked at the top in mathematics and science in both the TIMSS and PISA and were ranked near the top in reading literacy (Korea was ranked 6th and Japan was ranked 8th). These results baffle many psychologists because the results directly contradict existing psychological theories.

More than a century ago, psychological testing and assessment was initiated by Galton to measure innate ability. American psychologists (e.g., Terman at Stanford University, Yerkes at Harvard University, Garrett at Columbia University) continued the tradition. Terman developed the Stanford–Binet intelligence quotient (IQ)

test to measure inborn intelligence and to document individual, ethnic, racial, and sex differences. The results of the test were used to shape U.S. national policies (e.g., forced sterilization of "feebleminded" individuals, segregation of races, restricting Asian immigration). In 1924, Asians were considered a "genetically inferior" race, and the passage of the National Origins Act barred Asians from immigrating to the United States until 1965. Also, the IQ test scores were used to justify racial and sex segregation.

In 1977, Kamin stated that when the first version of the Stanford–Binet test was administered and published in 1916, one set of results contradicted the expectations in that girls of all ages outscored boys by an average of 2 to 4%. Terman unilaterally deleted, revised, or added new items so that this difference disappeared, and subsequently boys did better than girls. Thus, the IQ test was developed and revised not on a strict scientific basis but rather to fit the preconceptions at the time.

In 1960, Korea had one of the lowest literacy and educational achievement levels in the world. Currently, Korean students receive one of the highest scores in mathematics, science, and reading literacy. These dramatic reversals in sex and national differences cannot be attributed to genetic basis but can be attributed to social, economic, and cultural factors. Existing psychological theories argue that innate born differences (e.g., IQ, personality), ability attribution, intrinsic motivation, and self-esteem are important factors in predicting educational achievement. However, a series of studies conducted by Stevenson and colleagues and by Kim and colleagues challenged the basic assumptions of these theories.

The success of East Asian students can be traced to the strong sense of relatedness; social support provided by parents; belief in effort, hard work, and self-discipline; and coordinated efforts of parents, teachers, and governments. First, students, parents, teachers, and administrators agree that educational success is the most important goal for students, and they work together in achieving this outcome. Second, East Asians believe in effort, discipline, and persistence as a means to their goals. People believe that if they work hard and discipline themselves, they can accomplish their desired objectives. Third, the sacrifice and support provided by parents are essential ingredients in their children's educational success. Emotional support in the form of encouragement, praise, security, and understanding, rather than expert advice from professionals, is reported as an important factor for their success. Fourth, East Asians believe that ability can be acquired and personality can be polished through persistent effort rather than believing that these characteristics are innate.

When self-concept in mathematics and science was assessed, Korean students were ranked 32nd and 21st, respectively, whereas U.S. students were ranked 1st and 4th, respectively (TIMSS). Korean students did not believe that they were good in mathematics and believed that they needed to work hard to do well, and this belief and persistent effort led to the attainment of high scores. In contrast, American students believed that they were good in math and did not feel the need to study hard because expending effort implied a lack of natural ability. As a result, they did not do well on these tests. In terms of motivation, when Korean students were asked about their reasons for studying mathematics, 85% indicated that attending a desired university was among their reasons, 62% cited pleasing their parents, and 44% cited obtaining a desired job. In other words, their reasons for studying mathematics were largely extrinsic motivation.

Finally, using the indigenous psychology approach, a 6-year longitudinal study was conducted to examine the factors that influence academic achievement of Korean students. The role of parents emerged as the most important factor, followed by the role of teachers and the role of friends. The current authors found, in a structural equation model, that influence from parents (e.g., social support, achievement pressure, expectations) had a direct effect on students' achievement motivation, efficacy beliefs, and studying time and, in turn, had both a direct and an indirect effect on academic grades. Moreover, a feeling of indebtedness and respect toward their parents increased students' achievement motivation and efficacy beliefs and, in turn, increased the amount of time they studied and improved their academic grades. The socioeconomic status of the parents affected how much support they provided to their children. These results point to the role of external factors (e.g., socioeconomic status of parents, the role of significant others) and the belief in effort as being responsible for the phenomenal academic achievement of Korean students. Individual factors (e.g., self-esteem, intelligence, personality, intrinsic motivation, ability attribution) had very little impact on academic achievement.

6. FUTURE DIRECTIONS

Although science can provide the most accurate understanding of the world, it can also blind and limit one's understanding. Researchers start with an idea, model, theory, or method to uncover yet another secret of life. Researchers' preconceptions can both aid and limit

scientific discovery. As noted previously, psychologists' emphasis and focus on individualistic and intrinsic factors limited their understanding of external factors that are critical to students' academic achievement. A number of external impositions have limited the development of psychology, and indigenous psychology represents being freed of these impositions.

First, psychologists used the natural sciences model to study humans. In the rush to become a respected branch of science, early psychologists tailored the psychological phenomena to fit the natural sciences paradigm. Psychologists were able to achieve a modest degree of methodological sophistication, but central human qualities such as intentions, meaning, and agency have been omitted to distort one's understanding of psychological phenomena. Indigenous psychology advocates the development and use of appropriate methods and concepts to study psychological phenomena.

The second imposition is the assumption of the universality of psychological theories. With very little development, testing, and data, psychological theories are assumed to be universal. This assumption is particularly problematic because most theories were developed in the United States and tested mainly on university students in laboratory settings. In other words, theories that were tested on less than 1% of the total population have been assumed to be universal. Indigenous psychology emphasizes the documentation and analysis of psychological variations before conclusions can be made about their universality.

Third, expert or professional knowledge has been imposed on the lay public. Psychological theories and concepts have been used to explain and modify behavior. In most cases, the predictive value of these theories is very low compared with that of the natural sciences. Psychologists have largely failed to describe psychological phenomena from the inside, that is, from the experiencing persons. Instead, psychologists have dissected the world into abstract concepts and theories that are removed from the "real world" in which people live.

Indigenous psychology advocates being freed of these external impositions and obtaining the phenomenological information as insiders as well as through impartial third-person analysis. Indigenous psychology advocates linking knowledge from humanities (which focus on human experience and creativity) with knowledge from social sciences (which focus on analysis and verification). Most attention has been focused on internal or external validity and not on practical validity. In other words, do the theories help to understand, predict, and manage human behavior? It is known that most psychological theories (e.g., behaviorist, Freudian, Piagetian, humanist) do not withstand the test of empirical scrutiny. The main reason is that these theories have been developed using limited information (e.g., studying laboratory rats, interviewing too few patients, observing one's children), and they have not developed a process for systematic documentation, understanding, and verification of psychological phenomena.

The people who understand psychological phenomena well might not be psychologists. In fact, perhaps the greatest psychologist was Shakespeare. He was not an analyst like Freud or Piaget, and he did not conduct experiments like Skinner, but Shakespeare was able to capture human drama both on paper and on stage. His dramas have been performed for more than four centuries in many cultures and are loved throughout various parts of the world. In a similar vein, the greatest therapist might have been Beethoven or Mozart, whose music is able to soothe frazzled nerves and the frustrations of daily life. Disney could be considered as the most notable developmental psychologist because he was able to capture the hearts and minds of the young as well as the "young at heart." We might not think of these people as psychologists, but they captured and reproduced human psychology on stage, film, and paper. It is the role of indigenous psychologists to provide an understanding that is both analytical and phenomenological—that is scientific, practical, and relevant to people's daily lives.

See Also the Following Articles

Educational Achievement and Culture ■ Intrinsic Motivation and Self-Determination ■ Motivation and Culture

Further Reading

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Berry, J. W., Poortinga, Y. H., Segall, M. H., & Dasen, P. R. (1992). *Cross-cultural psychology: Research and applications*. Cambridge, MA: Cambridge University Press.

Chorover, S. L. (1980). *From genesis to genocide: The meaning of human nature and the power of behavior control*. Cambridge, MA: MIT Press.

Enriquez, V. G. (1993). Developing a Filipino psychology. In U. Kim, & J. W. Berry (Eds.), *Indigenous psychologies: Research and experience in cultural context* (pp. 152–169). Newbury Park, CA: Sage.

Kamin, L. J. (1977). *The science and politics of I.Q.* New York: Penguin Books.

Kim, U. (2001). Culture, science, and indigenous psychologies: An integrated analysis. In D. Matsumoto (Ed.), *Handbook of culture and psychology* (pp. 51–76). Oxford, UK: Oxford University Press.

Kim, U., & Berry, J. W. (1993). *Indigenous psychologies: Experience and research in cultural context*. Newbury Park, CA: Sage.

Kim, U., Helgesen, G., & Ahn, B. M. (2002). Democracy, trust, and political efficacy: Comparative analysis of Danish and Korean political culture. *Applied Psychology: An International Review*, *51*, 317–352.

Kim, U., & Park, Y. S. (2003). An indigenous analysis of success attribution: Comparison of Korean students and adults. In K. S. Yang, K. K. Hwang, P. Pedersen, & I. Daibo (Eds.), *Progress in Asian social psychology: Conceptual and empirical contributions* (pp. 171–195). New York: Praeger.

Kim, U., Park, Y. S., & Park, D. H. (1999). The Korean indigenous psychology approach: Theoretical considerations and empirical applications. *Applied Psychology: An International Review*, *45*, 55–73.

Park, Y. S., & Kim, U. (2000). The impact of changing parent–child relationships on adolescents' functioning: Comparison of primary, junior high, senior high, and university students. *Korean Journal of Educational Research*, *38*(2), 109–146.

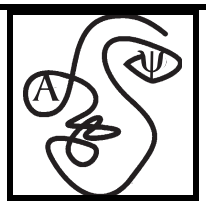
Park, Y. S., & Kim, U. (2003). Self-efficacy, achievement motivation, and academic achievement of Korean students: Indigenous psychological analysis. *Korean Journal of Educational Psychology*, *17*(1), 37–54.

Sinha, D. (1997). Indigenizing psychology. In J. W. Berry, Y. H. Poortinga, & J. Pandey (Eds.), *Handbook of cross-cultural psychology*, Vol. 1: *Theory and method* (pp. 131–169). Boston: Allyn & Bacon.

Stevenson, H., Azuma, H., & Hakuta, K. (Eds.). (1986). *Child development and education in Japan*. New York: Freeman.

Stevenson, H., & Lee, S. Y. (1990). Context of achievement: A study of American, Chinese, and Japanese children. *Monographs of the Society for Research in Child Development*, *55*, 1–75.

Wirth, L. (1946). Preface. In K. Mannheim, *Ideology and utopia: An introduction to sociology of knowledge*. New York: Harcourt, Brace.



Industrial/Organizational Psychology across Cultures

Zeynep Aycan
Koç University, Istanbul, Turkey

1. Introduction
2. The Role of Culture in Industrial and Organizational Psychology
3. Cross-Cultural Approaches to Work Motivation
4. Cross-Cultural Variations in Human Resource Management Practices
5. Managing Cultural Diversity
6. Culture and Work Attitudes
Further Reading

GLOSSARY

culture Prevailing and shared values, norms, assumptions, belief systems, and behavioral patterns that differentiate one human group from another.

human resource management practices The policies and practices involved in carrying out the human resource aspect of a management position, including recruiting, screening, training, rewarding, and appraising.

motivation Set of processes that arouse, direct, and maintain human behavior toward attaining some goal.

organizational diversity Workforce variety with respect to gender, culture, national origin, handicap, age, and religion.

work attitudes and behavior Clusters of feelings, beliefs, and behavioral intentions toward any aspect of work or work settings.

Industrial and organizational psychology (I–O) is one of the major applied fields of psychology that is concerned with behavior in work situations. Because of globalization, there is an increasing interest in the influence of culture on work behavior among scientists and practitioners. In this article, the cross-cultural perspective in I–O psychology is discussed in four main areas. First, the influence of culture on work motivation is examined. Second, human resource management practices in the areas of staffing and performance management are presented. Third, research on managing cultural diversity is reviewed with a focus on expatriate management, women's career development, and effective functioning of multicultural teams. Finally, cross-cultural research on work attitudes and behavior is summarized in the areas of job satisfaction, organizational commitment, organizational citizenship behavior, and organizational justice.

1. INTRODUCTION

Industrial and organizational psychology (I–O) was defined by Blum and Naylor as "the application or extension of psychological facts and principles to the problems concerning human beings operating within the context of business and industry." During recent years, the interest in cross-cultural perspectives of I–O psychology has

©2004 Elsevier Inc.
All rights reserved.

increased, due mainly to changing economic and social conditions such as workforce diversity, availability of telecommunications across cultures, and globalization and global competition. This article provides an overview of the cross-cultural I–O psychology literature in the areas of motivation, workforce diversity, human resource management (HRM), and work attitudes and behavior. Although leadership is an area that has attracted the most cross-cultural I–O research attention, it is not included in this article because a full treatment of the topic is available elsewhere in this encyclopedia.

2. THE ROLE OF CULTURE IN INDUSTRIAL AND ORGANIZATIONAL PSYCHOLOGY

Organizations are complex systems that operate under the influence of multiple environmental forces that are both internal (e.g., size, type of work, industry and production, type of workforce, technology, stage of development) and external (e.g., political, legal, educational, institutional) to the organization. The challenge of cross-cultural I–O psychology is to disentangle the impact of culture (i.e., sociocultural context) from other internal and external environmental forces.

Since the early 1960s, comparative studies of organizations have focused on noncultural factors influencing organizational structure and practices. Among the most popular noncultural approaches is the "contingency" approach, within which there were four main streams. The first, referred to as the "logic of industrialization," asserts that industrialization has a homogenizing effect on organizations around the world, irrespective of the political, economic, and cultural contexts. The second thesis, generally referred to as "technological implications," suggests that technological advancement and automation leads to a transformation of social relations and attitudes at work (e.g., more control over work schedule and work processes, increased emphasis on developing social networks). The "culture-free contingency theory of organizations" approach emphasizes the role of contextual elements such as size of the organization, industry, and dependence on other organizations. The final thesis emphasizes the role of strategic development of organizations, according to which organizations are transformed from small, less hierarchical, and domestic structures to large, complex, professional, and international structures. This transformation has implications for practices such as planning, diversification, and role differentiation.

Another type of noncultural approach stems from the political–economy perspective. In this perspective, organizations in the same sociopolitical systems (e.g., capitalism vs socialism) are assumed to have similar characteristics, especially with respect to organizational objectives, control strategies, and degree of centralization and decision making. The final noncultural approach, namely the "societal effect approach," takes into account the social context in which organizations operate, with specific emphasis on the educational system, the system of industrial relations, and the role of the state.

Critics of the noncultural approaches are concerned with the deterministic orientation of these approaches as well as their underestimation of the role of culture in explaining organizational phenomena. Some scholars take an interactionist perspective, suggesting that culture influences some aspects of organizational practices more than it does others. For instance, whereas organizational contingencies influence the "formal" characteristics of organizations (e.g., centralization, specialization, span of control), cultural variables influence the "interpersonal" aspects (e.g., power and authority structure, delegation, consultation, communication patterns) or the "organizational processes" (i.e., the way in which organizations function). Others assert that culture has a moderating effect on organizations; that is, even though the contingencies help to determine the organizational structure, culturally driven preferences influence the exercise of choice among alternative structures. For instance, the effect of industrialization on organizations is not homogenous in every country (e.g., the case of Japan). Similarly, within the capitalist system, there is wide variety among organizational and managerial practices. Researchers argue that similarities in socioeconomic systems resulted in similar organizational objectives, but the ways in which these objectives were materialized differed depending on the cultural contexts. The preceding discussion, therefore, underlines the important role that culture plays in understanding organizational phenomena. However, culture's effect is more salient in some aspects of organizations than in others. The following sections present the state of affairs in those areas of I–O psychology where the impact of culture is significant.

3. CROSS-CULTURAL APPROACHES TO WORK MOTIVATION

Work motivation is one of the most popular research topics in cross-cultural I–O psychology. The "content

theories" of motivation address the question of what motivates employees, whereas the "process theories" attempt to answer the question of how employees are motivated. Studies conducted in both paradigms clearly demonstrate the influence of culture on motivation.

3.1. Cross-Cultural Differences in What Motivates Employees at Work: Content Theories

Cross-cultural research on motivation between the 1970s and the 1990s focused mainly on the comparison of needs and values. The content theories of motivation assert that goal-directed behavior is initiated to satisfy unfulfilled needs, and values are considered the cognitive representation and transformation of needs. Early cross-cultural research tested the validity of various need taxonomies, including those of Maslow, Alderfer, and McClelland. Employees in economically developed and developing countries were found to be motivated by different incentives due to variations in needs. For example, malnutrition was the main cause of low motivation and productivity in Nigeria, whereas job security was most needed, and hence most motivating, in India and Guatemala. Similarly, satisfaction of physiological and safety needs were more salient for cane cutters, whereas affection and esteem needs were most salient for fishermen.

In subsequent studies, it was shown that in collectivist societies, "self-actualization" (i.e., the highest order need) is replaced by "service to the in-group" or actualizing the in-group's potential (i.e., "collectualization") as the highest ideal to motivate workers. Also in collectivist cultures, "achievement" has collectivist overtones such as respect for others, fulfilling collective goals, and glorifying established virtues. The conclusion of the cross-cultural research on need-based theories of motivation is that the structure of needs may be universal (e.g., lower order needs such as safety, higher order needs such as esteem) but that the relative strength or the contents of the standards vary by culture.

The other strand of research under the content theories emphasizes the importance of "values" as a motivating force. A major research project was the Meaning of Work (MOW) studies by the MOW International Research Team that was carried out in the United States, Japan, and six Western European countries. The MOW team was primarily interested in the extent to which work is important in people's lives in various cultural contexts. Their assumption was that work is more motivating if it is central to one's life. Results showed that work was most central in Japan and least central in West Germany (with the United States in the middle). More recent studies replicated these findings; Japanese employees are motivated most by the task itself, whereas their U.S. counterparts are motivated mostly by affiliation and recognition.

There are other large-scale cross-cultural studies on the motivating value of work. One was conducted by the Work Importance Study (WIS) group in 11 diverse countries, including Croatia, South Africa, and Portugal. The findings showed significant cross-cultural differences with respect to the extent to which work is a central life value compared with family, community, and leisure. Another study, called the Work Socialization of Youth (WOSY) project, examined the patterns of meanings of work from a longitudinal perspective in 7 West European countries and Israel. The results demonstrated that the meaning of work was changing across cultures and over time. Finally, the Chinese Culture Connection team demonstrated the salience of the "Confucian work ethic" emphasizing moral duty and obligation to the collectivity as being the primary motivating factors.

3.2. Cross-Cultural Differences in How Employees Are Motivated at Work: Process Theories

Process theories deal with the ways in which employees are motivated at work. There are four major process theories of motivation: goal-setting theory, job enrichment theory, equity theory, and expectancy theory. Among the four, goal-setting theory of Locke and Latham has received the most attention from cross-cultural researchers. The theory asserts that the existence of goals is motivating and that goals with particular characteristics (e.g., challenging, specific, accepted) are more motivating than others. Participation and joint decision making are considered to be a key factor in motivation. Participation enhances motivation in that it provides greater autonomy and control over work outcomes and reduces anxiety and uncertainties. The Industrial Democracy in Europe (IDE) International Research Group compared the participative decision-making practices in 12 European countries in 1981 and repeated the study 10 years later in 10 of the original 12 countries as well as in Japan and Poland. This was a major undertaking to show that participation had different meanings and practices across cultures. These

studies show that in the most collectivist cultural contexts, group participation is more motivating and effective in improving work outcomes than is individual participation. Individual and face-to-face involvement in decision making (i.e., the North American model) is, in fact, not desired in all cultural contexts. In cultures that Hofstede identified as high in power distance and uncertainty avoidance, participation may even result in demotivation because subordinates seek guidance from their superiors to reduce uncertainties. Research suggests that for participative goal setting to be motivating in developing countries, organizations should enhance employees' self-efficacy beliefs through empowerment, coaching, and mentoring.

The job enrichment theory of Hackman and Oldham suggests that workers are motivated to the extent that they experience meaningfulness in their jobs. This can be achieved through more autonomy and discretion in one's job, feedback on performance, opportunity to use multiple skills (i.e., skill variety), identification of the contribution of one's job to the overall work context (i.e., task identity), and knowledge of the significant contribution of one's job to others' lives (i.e., task significance). Jobs that are designed to have high autonomy, feedback, skill variety, task identity, and task significance are referred to as "enriched jobs." This theory is based on individualist assumptions of employees' desire for independence, freedom, and challenge. Thus, it is argued that job enrichment might not be as effective in countries with power hierarchies and avoidance of uncertainties. Increasing discretion and responsibility elevate anxiety and uncertainty, and worker autonomy defies the role of the authority. In more collectivist European countries, job design takes the form of "autonomous work groups" in which job enrichment principles are implemented at the group level (e.g., team autonomy, team-based feedback). However, this system is reported to have limitations in developing countries due to uncertainty avoidance and resource scarcity that is endemic to the majority of developing countries.

The equity theory of Adams purports that employee motivation is achieved by invoking the perception of fairness. One way in which to increase fairness is to establish an equitable relationship between "inputs" and "outputs" among workers. Inputs represent everything that employees perceive as their contribution to the job (e.g., training, skills, time, effort, loyalty, commitment to the organization), whereas outputs represent everything that employees perceive as received in return for their contribution (e.g., pay, promotion, praise, recognition and awards). Equity theory has not attracted much cross-cultural research attention, but the available evidence suggests that what constitutes inputs and outputs varies across cultures. In collectivist cultures, loyalty, commitment, seniority, and tenure are more important inputs for promotion and rewards than is job performance. Also, in cultures where there are status hierarchies, the social class or caste to which one belongs is considered an input. Similarly, there are cultural variations in the perception of outcomes. Challenge and autonomy in the job and salary increases are considered to be valued outcomes in individualist cultures, whereas praise from supervisors can be an important outcome in collectivist cultures. There are ways in which to restore equity such as changing inputs (e.g., working less), changing the comparison group, changing the outcomes (e.g., asking for a salary increase), and leaving the organization. There is not enough research to indicate cross-cultural variations in the ways in which to restore equity. However, there is a large body of cross-cultural research in the area of organizational justice. Reward allocation based on an equity principle is not endorsed in all cultural contexts.

The final process theory of motivation is the expectancy theory of Vroom. According to the theory, employees are motivated to the extent that their expectations are met in the following ways. First, if they exert enough effort, their job performance will be at the desired level. Second, if they perform at the desired level, it will lead to some outcomes. Third, the outcomes will have high valence or attractiveness. Although there is no systematic research testing the validity of the theory across cultures, it seems to have more applicability in individualist cultures than in collectivist ones. In individualist cultures, employees are motivated if there is a high likelihood of meeting their own expectations, whereas in collectivist cultures, motivation is related to the probability of meeting the expectations of "significant others" (e.g., family, organization, superiors). "Face saving" is also an important motivator in collectivist cultures. Another cross-cultural variation may exist in people's beliefs in the controllability of events. In fatalistic cultures, employees do not believe that control of events is in their hands; that is, they have an "external locus of control." Thus, expectancy theory might not be useful in such cultural contexts.

3.3. The Culture-Based Model of Work Motivation

The first theoretical model explaining how culture influences motivational practices and their outcomes was

proposed by Erez and Earley in 1993. According to the cultural self-representation model, motivation is related to the employee's self-concept. Cultural context determines how one perceives oneself. To the extent that motivational practices are congruent with culturally based self-construals (e.g., the way in which the self is defined and perceived), there is a high likelihood of performance and satisfaction. For example, if in teamwork employees receive rewards based on their individual contributions to the team, this is congruent with the cultural context where the self is perceived as an independent entity (i.e., individualism), whereas group reward is more motivating in the cultural contexts in which employees have interdependent self-construals (i.e., collectivism). The congruence results in three self-derived motives: self-enhancement (maintaining a positive cognitive and affective state about the self), self-efficacy (perceiving oneself as competent and efficacious), and self-consistency (experiencing coherence and continuity). For employees who have interdependent self-construals, individual-based rewards in teamwork lower the positive self-representation and hence motivation. In general, employees are motivated through organizational practices that maintain and improve their self-enhancement, self-efficacy, and self-consistency.

4. CROSS-CULTURAL VARIATIONS IN HUMAN RESOURCE MANAGEMENT PRACTICES

With increasing globalization, staffing and performance management have become the keys to success of multinational firms. In the HRM field, these are the two topics that have received more cross-cultural research attention than have other areas of HRM (e.g., human resource planning, job analysis, compensation, reward management).

4.1. Staffing: Recruitment and Selection

Culture influences recruitment and selection in many ways such as in attitudes toward selection and testing, the purpose that staffing serves, and the perceived fairness and appropriateness of criteria and methods that are used in the process. The first theme of cross-cultural variation is the attitudes toward selection and testing. In the North American context, selection is a one-way process whereby applicants are tested to predict their future work behavior. However, in the European context, selection is a mutual agreement and negotiation between the organization and the candidate. In some European countries such as Italy, France, Sweden, and Portugal, the issue of testing has a negative connotation because it is perceived as invasion of privacy, violating individuals' rights to control their own careers and creating a barrier to the holistic representation of oneself.

The main purpose of recruitment and selection also varies across cultures. In North America, strategic human resource planning is geared toward recruiting the right number of people with the best qualifications to do the job. In contrast, in India and Eastern Europe, organizations hire more employees than are needed so as to combat poverty and unemployment. In such cultural contexts, the strategic goal of organizations is to meet societal needs as well as economic ones.

The criteria for selection should be perceived as relevant to the job and as predictive of future performance. In the United States, some of the most commonly used criteria for selection include education, cognitive skills, job experiences, and personality traits. There is a paucity of cross-cultural research on the appropriateness of such selection criteria in other cultural contexts. However, the few existing studies cast doubt on the appropriateness of criteria such as education. It is argued that knowledge and skills acquired by formal education might not be the best predictors of performance in countries where education either is not available to everyone or is inadequate in preparing individuals for the workforce. A few studies suggest that criteria for selection are more interpersonal than individual in collectivist cultures. For example, in collectivist cultures, social selection criteria are used—team members' favorable opinions about the candidate in Japan; agreeableness, good interpersonal relations, and trustworthiness in Islamic Arab countries; belonging to the same in-group as the manager (e.g., the same family or homeland) in India; and positive attitudes toward family life in Latin America.

Once criteria are established, the next step involves determining the method of recruitment and selection. Fairness is related to the candidates' perception of the appropriateness of recruitment channels, the questions that are asked on application forms, and job previews (i.e., a brief and a realistic review of the job context and the work environment). Word-of-mouth is a common method of announcing a job opening in cultures where the hiring of in-group members is favored. Some questions on the job application forms (e.g., marital status, religious and family backgrounds, physical appearance)

could be perceived as offensive in some cultures more than in others. In job previews, a detailed description of a specific job position will not be a requirement of an ideal recruitment process in cultures where "generalists" are preferred over "specialists" (e.g., Japan).

The methods of selection used in various countries have attracted some cross-cultural research attention, but countries included in these studies have been limited mainly to those in North America and Northwestern Europe. In these studies, biodata was found to be among the least common methods of selection, whereas the interview was among the most common. In a majority of countries, references or recommendations are used for different reasons and to varying degrees. For instance, in the United Kingdom, the United States, and Australia, recommendations are used only as a final check, whereas they are relied on heavily in Southeastern European countries and in India.

In the North American literature, studies on the predictive validity of various selection methods have increased in number over recent years. However, the social validity (i.e., perceived fairness and acceptability of a particular practice) is an issue that is overlooked in the cross-cultural literature. For instance, in one study, French job applicants perceived written ability tests to be less impersonal and perceived personality tests to be more offensive in violating privacy than did American applicants.

This brief review of the literature shows that there are cross-cultural variations in various aspects of the recruitment and selection process.

4.2. Performance Management: Criteria, Evaluation, and Development

Criterion development and measurement has been an issue of ongoing debate among I–O psychologists. There is neither a uniform description of "performance" nor consensus on the dimensions of it. The criterion problem is exacerbated at the cross-cultural level because both the definition and structure of "good performance" are culture bound. According to some of the recent North American theoretical formulations, the most salient dimensions of performance are job-specific knowledge and proficiency, communication competence, effort, mentorship/supervision, quality of output, productivity, interpersonal competence/teamwork facilitation, and administrative competence. Cross-cultural replications of these dimensions have yet to be done.

What constitutes good performance in various cultural contexts? Performance-orientated cultures place more emphasis on the end result rather than on the process. As such, performance criteria are objective and measurable. In collectivist and high-power distance cultures, employee loyalty is valued more than productivity. Outcomes are important, but social and relational criteria, which are more subjective, are weighted heavily in evaluating employees. Such criteria include good human nature, harmony in interpersonal relations, trustworthiness, respectful attitude, loyalty and deference toward superiors, effort and willingness to work, awareness of duties and obligations, gratitude, organizational citizenship, conformity, and contribution to team maintenance.

Unlike individualist cultures, collectivist cultures downplay individual differences. Therefore, the primary purpose of performance evaluation is not differentiating among individuals but rather justifying decisions on compensation and promotion. Some authors purport that appraisal is used to control and instill loyalty. Formal evaluations are conducted less frequently (because the primary purpose is not to provide feedback), and their timing coincides with salary adjustment and promotion periods. In high-power distance cultures, performance evaluations may also serve the purpose of reinforcing the authority structure. In such cultural contexts, performance is usually evaluated by superiors (i.e., immediate supervisors, managers, and/or several levels up). Because performance appraisal is a top-down unilateral process, evaluation of superiors is rare in high-power distance cultures. Therefore, the 360-degree performance appraisal method (i.e., self-evaluation plus evaluation by superiors, subordinates, peers, and customers) that is popular in the United States is not appropriate in high-power distance cultures. In collectivist cultures, the 360-degree evaluation may disturb group harmony due to constant monitoring of the behavior of one's colleagues. The accuracy of such evaluations is also jeopardized because in-group favoritism and loyalty prevent assignment of low performance ratings to in-group colleagues and superiors.

Rating errors are barriers to obtaining reliable performance evaluations. The most pervasive ones include attribution and leniency errors. The accuracy of the process depends partly on the correctness of attribution of success and failure. Cross-cultural variations in attribution are well documented. For instance, Asians usually attribute their success to luck and help received from others, whereas they usually attribute failure to their lack of ability and inadequate personal effort. In a collectivist cultural context, those who

engage in self-effacement gain acceptance by peers, whereas those who take pride in individual achievements receive disapproval. Modesty in self-evaluations leads to negative leniency bias, that is, assigning a lower evaluation to oneself than is deserved.

The final topic is related to the communication of performance feedback. Culture has a bearing on the way in which feedback is given and received. Regardless of the cultural context, performance feedback is not easy to give or receive, especially when the feedback is of a negative nature. However, in some cultures where the distinction between life and work space is blurred, feedback on one's job performance is perceived as equivalent to a rejection of the person, who then experiences loss of face. On the other hand, positive feedback on performance is also not common in collectivist cultures. Positive feedback on individual performance could disturb group harmony because it may induce jealousy and resentment among those who did not receive such feedback. Also, in collectivist cultures, positive feedback is expected to come from the outside. If a manager praises his or her own employees, it could be perceived as inappropriate and self-serving.

If neither negative nor positive feedback is appreciated, this implies that no feedback could be the most preferred condition in collectivist cultures. There is at least one empirical study that supports this view, showing that Japanese and Chinese employees do not take any initiative in seeking feedback on individual performance. This is explained by three factors. First, as mentioned previously, there is avoidance of negative or positive feedback. Second, seeking feedback on individual performance, rather than on group performance, is perceived as "self-centeredness." Finally, communication is indirect and subtle in collectivist cultures. Performance-related information is usually embedded in contextual cues rather than revealed in a direct and confrontational manner.

The final phase in performance management concerns the improvement of performance. The most frequently administered practices used to improve performance include performance-based rewards, training, and coaching/mentoring. Are rewards contingent on performance? The answer depends on the cultural context. It was found that in collectivist and fatalistic cultures, there is a weak performance–reward contingency due to a number of factors. First, in fatalistic cultures, managers believe that employees are not malleable; therefore, managers do not see any point in rewarding employees as a means of changing work behavior. Second, individualist cultures use the norm of equity in allocating rewards. Therefore, there is one-to-one correspondence between performance and reward in individualist cultures, whereas reward allocation in collectivist cultures depends on contingencies other than performance. Finally, in collectivist cultures, rewarding the group as a whole is preferred over rewarding individual members. As mentioned previously, individual gratification may result in disturbed group harmony and resentment.

The second practice used to enhance performance is training. One of the first questions to ask is whether or not training needs are determined on the basis of performance outcomes. This seems to be the case in individualist and performance-oriented cultures. In low-performance-oriented and high-power distance contexts, decisions about who will participate in trainings are based on criteria other than job performance. Employees who maintain good relations with higher management are sent to desirable training programs (i.e., training overseas or at resorts) as a reward for their loyalty. Because training is given as a favor or reward, trainees provide highly favorable feedback on trainings and introduce a bias to the evaluation of training effectiveness.

The next practice in performance development is coaching and mentoring, which involves a personalized subordinate–superior relationship to enhance performance. Mentors are senior and more experienced people who provide guidance to employees. The role of a mentor includes modeling correct behavior, setting clear performance standards and goals with employees, coaching, monitoring and giving constructive feedback on performance, and making necessary adjustments to tasks and work environments to improve performance. Mentors work with and for employees. High power distance is not compatible with mentorship. In high-power distance cultures, superior–subordinate relationships are not personalized and managers are not usually participative in goal setting. However, mentorship is more likely when high power distance is combined with paternalism. In paternalistic cultures, superior–subordinate relationships are personalized and managers guide employees in professional as well as personal matters.

4.3. The Culture-Based Model of Human Resource Management

Kanungo and colleagues proposed and tested a theoretical model explicitly linking culture to HRM practices.

The model of culture fit asserts that variations in HRM practices are due to the internal work culture, which in turn is influenced by both the sociocultural context and the enterprise environment. The internal work culture is composed of managerial beliefs and assumptions about two fundamental organizational elements: the task and the employees. Managerial assumptions pertaining to the task deal with the nature of the task and how it should be accomplished best, whereas those pertaining to the employees deal with employee nature and behavior.

Organizations design and implement HRM practices based on the prevailing managerial assumptions about the nature of both the task and the employees. These assumptions, in turn, are influenced by two major environmental forces. First, the enterprise environment includes the ownership status (private sector vs public sector), industry, market competitiveness, and resource availability. Second, employee-related assumptions are influenced by the characteristics of the sociocultural environment.

The model of culture fit was tested in various cultural contexts, including Canada, the United States, Romania, Germany, Israel, India, Pakistan, China, Turkey, and Russia, with the following key findings. In fatalistic cultures, managers assume that it is not possible to change employee nature and behavior. Accordingly, they do not provide feedback or try to improve performance through rewards. In cultures where there are power hierarchies, managers do not hold the assumption that employees are proactive; rather, they assume that employees are reactive (i.e., expecting close guidance and supervision). As a result, managers do not give autonomy and discretion to employees in their job assignments. In collectivist cultures, managers assume that organizations and employees should fulfill obligations to one another. Based on this assumption, they provide empowering supervision both to fulfill their obligations toward workers and to enable workers to fulfill their obligations toward others in the organization. The model of culture fit is an important initial step toward understanding how, and to what extent, culture influences HRM practices. Cross-cultural and multilevel testing of the model is continuing.

5. MANAGING CULTURAL DIVERSITY

This section examines cultural diversity under three categories: expatriate management, women's career advancement, and multicultural teams.

5.1. Expatriate Management

Expatriates are employees of business or government organizations who are sent by their organizations to related units in countries other than their own for predesignated temporary time periods of usually more than 6 months and less than 5 years. In global competition, successful expatriation not only gives multinational companies a competitive edge, it also contributes substantially to professional growth of expatriates. Despite the importance attached to international assignments, failure to complete the full cycle of such an assignment and failure to work effectively on the job remain major problems. Criteria for effective expatriation include both adjustment to the new cultural context and high job performance. Difficulties in adjustment and performance result in high turnover rates and, therefore, failure to receive a return on investment. Stimulated by the observed problems in expatriation and the costs associated with it, this topic has emerged as one of the most popular topics for cross-cultural researchers.

Studies have identified a number of key factors influencing expatriates' cross-cultural adjustment and performance. The individual factors include those related to expatriates and their families. There is some research suggesting that personality characteristics, such as openness to new experiences, agreeableness, and tolerance for ambiguity, predict better adjustment outcomes. Competence in interpersonal relationships and willingness to learn about a new culture are factors that facilitate both adjustment and performance. Recent research also suggests that those individuals who have had cross-cultural interactions during their childhoods stand a better chance of succeeding in international assignments. Expatriate-related factors also include family adjustment issues. Expatriates whose families have adjustment difficulties are more likely to leave their assignments prematurely.

Organization-related factors include the preparatory stage and the HRM practices of the company. The majority of expatriates who fail in their assignments have been selected solely on the basis on technical and managerial competencies. Their adjustment potential and interpersonal competencies are usually not assessed. Similarly, many organizations fail to provide cross-cultural and country-specific training prior to departure. These problems are exacerbated by the uncertainties related to the assignment terms and job specifications. Difficulty in finding positions in the parent company on repatriation also lowers job motivation and performance after the employees return. Furthermore, the receiving

organization (i.e., the local unit) is usually ill prepared to manage the cultural diversity resulting from expatriates' arrival. Local units that offer training on multicultural team management provide better support and assistance to the incoming expatriates, thereby facilitating their adjustment and performance. Finally, cultural factors influence expatriation outcomes. In assignments to countries with large cultural distance (i.e., dissimilarity between the culture of origin and the culture of the host country), expatriates experience more difficulties.

5.2. Women's Career Development

According to the reports of international organizations (e.g., United Nations' *Human Development Reports, World Bank Reports,* International Labor Organization reports), women receive less pay, less occupational status, and less career advancement all around the world. Although women share similar experiences at work and in family, cross-cultural differences contribute to the magnitude of discrimination and conflict between work and family. The literature canvasses various approaches to explain women's difficulty in gaining equal access to career development opportunities. First, there is compelling evidence suggesting the effect of gender role stereotypes. In all cultures, women's primary responsibility is toward the family. In some cultures, women feel the freedom to choose between pursuing a career and raising a family; however, in the majority of countries, women do not have such a choice. Cross-cultural differences exist with respect to norms in childbearing. Career women may opt to have children in some economically developed countries (e.g., France), whereas this is not an option for women in many traditional cultures of developing countries. Moreover, in many collectivist cultures, women's family responsibilities include taking care of extended family members, in-laws, and aging parents, adding to women's total workload.

A second difficulty stems from the attitudes toward working women. In the majority of countries in the world, organizations hold negative attitudes toward women's career development. A common belief is that women are cognitively and psychologically ill prepared to become leaders in the business world. This belief is derived from socialization values and practices. In many traditional cultures, women are socialized to be shy, unassertive, dependent, passive, and obedient to the men in their lives. For example, according to the traditional beliefs in China, it is a virtue for a woman to have no leadership ability. Women themselves internalize these norms and develop unfavorable attitudes toward their career development.

Women are also cut off from male networks. As a result, they miss opportunities that would help them to advance in their careers. For instance, in Japan and Taiwan, evening gatherings in pubs or clubs are very common. However, because women's participation in such social gatherings is considered inappropriate, they are inadvertently cut off from the critical social (and hence informational) networks. Another negative attitude stems from the concern about women's and men's changing status in the family. In traditional societies, if a woman works, it is perceived to be the failure of the husband to provide for the family. It is also feared that women's economic independence will decrease men's power in the family.

Finally, women receive insufficient support, especially with respect to child care, from their families and the organizations for which they work. One of the most important sources of support is the spouse. In traditional societies, a man's helping out his wife with her family duties is uncommon. Husbands who do housework are stigmatized as being "woman-like," "submissive," "passive," or simply "weak." In such cases, the values and behaviors of husbands, especially those from high educational backgrounds, oscillate between traditional and modern. Support is also rare in organizations. Family-friendly work schedules and day care facilities are rare in many countries. Although managers are more tolerant and understanding of women's family responsibilities in paternalistic and collectivist cultures, such tolerance does not indicate support for women's career advancement.

5.3. Multicultural Teams

Teamwork is one of the key factors for organizational success. With increasing globalization, effective management of multicultural teams has become a critical issue. Research in this area is still developing, and conclusive findings are few. There are mixed findings regarding the performance of culturally heterogeneous teams compared with that of homogenous ones. There are four factors that influence the functioning of multicultural teams. First, differences in core values and norms regarding how members should behave and interact with one another lower the performance at initial stages. For example, members from a collectivist culture value interpersonal harmony more than they do task accomplishment. As such, they avoid confrontations and giving negative performance feedback.

There is at least one study indicating that social loafing (i.e., the tendency to withhold effort when performing a group task) occurs less in teams in collectivist cultures than in teams in individualist cultures. In collectivist cultures, team members work hard to achieve the group goals, and members feel obligated to contribute to the team. Members of high-power distance cultures prefer teams with leaders who guide them closely, whereas members of low-power distance cultures may function better in autonomous or leaderless teams. Differences in norms, values, and preferences are potential barriers to effective functioning in multicultural teams.

The second factor that influences team performance is the perceived cultural distance of the team members. Recent research found a curvilinear (i.e., inverted U-shaped) relationship between the perceived cultural distance and team performance in such a way that culturally homogenous and highly heterogeneous teams outperform moderately heterogeneous teams. Formation of subgroups is more likely to occur in moderately heterogeneous teams than in highly diverse teams. This lowers the team coherence and performance.

The third factor is time. Performance in diverse teams is a function of time. Heterogeneous teams are found to outperform homogenous teams over time. It takes more time for diverse teams to build trust and shared values and norms. Initial inefficiencies (i.e., process losses) due to inadequate communication and feedback may be overcome in time.

Finally, task structure affects performance of multicultural teams. In jobs that require an in-depth understanding of a particular cultural context (e.g., launching a new product in a new cultural context), cultural knowledge is task related and, hence, increases team performance. Multicultural teams perform better at tasks requiring idea generation and creativity than they do at production tasks requiring little discretion. Diverse teams function better under structured and noncompetitive task environments. Superordinate goals (i.e., goals that group members cannot reach without the help of the other group members) also facilitate team performance in multicultural teams.

Teams working in virtual environments (e.g., the Internet, videoconferencing) experience further difficulties in adjusting to uncertainties arising from technological problems. Virtual teamwork is more difficult for members who prefer face-to-face interactions (e.g., members of collectivist cultures) as well as contextual cues in communication (e.g., members of high-context cultures).

6. CULTURE AND WORK ATTITUDES

Work attitudes focus on how employees feel about their jobs and the organization. Exploring the cultural factors influencing work attitudes is one of the most popular research topics in cross-cultural I–O psychology because work attitudes are related to various work outcomes, including performance, productivity, turnover, and absenteeism. This section discusses the following cross-cultural differences in work attitudes: job satisfaction, organizational commitment (OC) and organizational citizenship behavior (OCB), and organizational justice.

6.1. Job Satisfaction

Work conditions (e.g., interpersonal relationships, pay, benefits, developmental opportunities) that are satisfying in one cultural context might not be satisfying in another cultural context. Research evidence suggests that workers in individualist cultures report more job satisfaction than do those in collectivist cultures. The higher level of job satisfaction in individualist countries is attributed to the fact that workers enjoy more individual freedom and rights and are more likely to fulfill their self-actualization needs. It has also been shown that people living in individualist countries report more life satisfaction, and this has a "spillover" effect in that it increases job satisfaction. However, recent research suggests that collectivism, rather than individualism, is positively correlated with job satisfaction. First, it is argued that employees "put up with" unpleasant work conditions because they believe that their work group members also suffer from the same conditions. The sense of sharing of this "common destiny" reduces dissatisfaction. Second, positive interpersonal relationships at work (e.g., collaboration, nonconfrontation) create a pleasant work atmosphere and enable workers to better cope with unpleasant conditions.

These seemingly conflicting findings are reconciled by research that identified contextual factors influencing job satisfaction. For instance, intrinsic job characteristics (e.g., opportunity for development, use of multiple skills, challenging job, autonomy) are more satisfying to employees in individualist cultures, whereas extrinsic job characteristics (e.g., pay, benefits, sense of security) and interpersonal aspects of the job are more satisfying to those in collectivist cultures. This variation is meaningful because job satisfaction is closely related to the

fulfillment of needs. In collectivist countries, people need to satisfy primarily their basic and interpersonal needs. As such, job characteristics such as challenge and autonomy do not yield job satisfaction unless salient needs are fulfilled. Furthermore, intrinsic job characteristics are positively associated with job satisfaction in countries with high social security (e.g., welfare system, national wealth) and small power distance, whereas this relationship does not exist in countries with low social security and large power distance. Employees in high-power distance cultures do not expect to participate in decision making, receive performance feedback, or be autonomous in their jobs. Lack of such intrinsic job characteristics does not result in dissatisfaction.

6.2. Organizational Commitment and Organizational Citizenship Behavior

OC is conceptualized as the extent to which employees feel a sense of allegiance toward their employers. The growing body of literature on OC suggests that it is an important organizational attitude predicting performance indicators (e.g., absenteeism, turnover intention), employee reactions to work, and employee well-being at work. In 1990, Allen and Meyer proposed a model of OC that consists of three components: affective commitment (i.e., employees' emotional attachment to and identification with the organization), continuance commitment (i.e., the perceived necessity to stay with the organization due to the high cost of leaving it), and normative commitment (i.e., the belief that one is obligated to stay loyal and committed to the employer).

Cross-cultural studies of OC have focused on two issues: structure of the construct (i.e., whether or not the three-factor model replicates across cultures) and correlates of OC. There is mixed support for the generality of the factor structure of the North American OC measures, and this casts doubt on the "universal" meaning of OC in various cultural contexts. The psychometric properties of these measures did not completely replicate in countries that are characterized by high collectivism, high power distance, and strong uncertainty avoidance (e.g., Korea). In such cultural contexts, the measurement holds better with the addition of some emic (i.e., culturally relevant) items. Cross-cultural variations are also found with respect to the correlates of OC. Antecedents and consequences of OC in Western European countries (e.g., Germany) were very similar to those in North America. Affective commitment was the best predictor of turnover intention in individualist cultures with high employment mobility, whereas normative commitment was the best predictor in collectivist cultures emphasizing long-term employment. Desirable organizational outcomes (e.g., low absenteeism, job satisfaction, employee well-being) are strongly correlated with affective commitment in North America but are linked to normative commitment in collectivist, high-power distance, and paternalistic cultures where loyalty is valued and rewarded.

OCB is described as the contribution of employees to the welfare and effectiveness of the organization by engaging in behaviors that are neither required nor expected. These behaviors are referred to as the "good soldier" or "extra-role" behaviors. Similar to the case of OC, cross-cultural research on OCB addresses the issues of whether or not its meaning and structure transcend culture. The original theory, as proposed in North America by Organ in 1988, describes five dimensions of OCB: altruism (e.g., voluntarily offering help to others), conscientiousness (e.g., being punctual, following company rules and regulations), courtesy (e.g., being mindful and respectful of others' rights), civic virtue (e.g., participating in the political life, attending meetings, speaking up on issues), and sportsmanship (e.g., not complaining, not nursing petty grievances, not gossiping, not falsely magnifying problems). Limited cross-cultural research in this area suggests that some of these behaviors, as described in the North American literature, are not appropriate in some cultural contexts (e.g., speaking up in high-power distance cultures), whereas others are not considered as in-role rather than extra-role behaviors (e.g., voluntarily offering help in collectivist cultures). Some evidence suggests that there are emic dimensions of OCB in collectivist cultures such as protecting company resources. Because of such variations, it is tentatively concluded that the dimensionality (e.g., the factor structure) of the Western OCB construct and its measures do not transcend culture. Similarly, there are variations in the predictors and outcomes of OCB. OCB is found to be predicted by the organizational justice perception in Western literature, whereas collectivist values are seen as better predictors of OCB in Eastern traditional cultures. OCB was correlated more strongly with turnover intention and productivity in individualist cultural contexts rather than in collectivist ones.

6.3. Organizational Justice

Organizational justice is concerned with the perception of fair treatment of employees in organizations. Justice perception has important implications for organizations

because it relates to OC, productivity, trust, and compliance with decisions. For more than a century, philosophers have been debating the concept of fairness with little agreement. Disputes over this issue are magnified when cross-cultural differences in values and norms regarding justice are involved in research. Despite such challenges, organizational justice has proven to be a fruitful area that has attracted a great deal of cross-cultural I–O research attention.

The literature draws a distinction between two types of organizational justice. The first is referred to as "distributive justice" and is concerned with the fairness of norms or criteria used in distributing resources. The second is "procedural justice" and refers to the fairness of the means (e.g., procedures, practices) used to achieve the organizational results.

Earlier cross-cultural research on distributive justice was limited to examination of the criteria used in performance appraisal. With respect to procedural justice, the focus was on the fairness perception of various leadership styles (e.g., participative leadership). The overarching assumption was that criteria influencing justice perception had similar meanings across cultures, but the cross-cultural differences consisted of different weights being given in different cultures. Social and organizational psychology literature during the 1980s offered an "interactionist" perspective purporting that the social context cued the application of fairness criteria. Such sophisticated models advanced the literature in such a way that some authors claimed that cross-cultural research had a significant influence on the mainstream (mainly North American) theories in the area of justice.

In the area of distributive justice, cross-cultural research has investigated the differential norms that are used in distributing resources. Resources are distributed on the basis of three norms: equity, equality, and need. The equity principle uses the input–output ratio as a method of allocating resources to employees, the equality principle advocates similarity in resource allocation regardless of the differences in inputs, and the need principle promotes the distribution of resources on the basis of need. Earlier cross-cultural research found that employees in individualist cultures perceive justice as being done when the equity norm is used, whereas those in collectivist cultures favor use of the equality and need norms. However, later studies showed that collectivist cultures prefer different norms depending on the context.

In 1997, Leung proposed a model describing situational variables interacting with individualism–collectivism in cueing differences in distribution criteria. According to the model, goal-directed behavior is the most immediate predictor of selection of the distributive criterion. For instance, the equity principle is perceived to be more appropriate in collectivist cultures when applied to out-group members if the salient goal is to increase productivity, whereas the need principle is favored when the immediate goal is to protect the well-being of others under the condition of resource scarcity. In applying the equity principle, what constitutes inputs and outcomes is culturally determined, but there is surprisingly little research on this topic. Limited available research suggests that in collectivist cultures, loyalty, family size, and networks (e.g., *guanxi)* are considered to be appropriate inputs. The criteria to be used also depend on power distance. In hierarchical societies, in-group favoritism is perceived to be fair in distributing resources. Moreover, in interpersonal comparisons, it is accepted that the more powerful ones are entitled to receive more benefits than are those at the bottom of the hierarchy. Researchers urge caution regarding changes in the perception of power, and hence the perception of justice, over time.

Cross-cultural research in procedural and interactional justice is very limited, but available evidence suggests that the perception of fairness in procedures and interpersonal treatment depends largely on cultural norms. For example, in individualist cultures, having a feeling of control over decisions increases the perception of justice, and this finding is also replicated in some collectivist cultures. Similarly, voice (i.e., the opportunity to express one's opinions) is perceived to be fair in individualist cultures, and this perception is mediated by one's relationship with the people in authority. In collectivist cultures, procedures that follow tradition are accepted as fair, suggesting that norms have a stronger impact on the perception of justice in collectivist cultures than in individualist ones.

There are three factors that increase the perception of procedural fairness. The first is the benevolent intent of the authority. Cross-cultural variations exist in the perception of the intent underlying the behavior of superiors. Some research suggests that workers in high-power distance cultures are more likely to tolerate harsh treatment by their superiors, partly because this behavior is perceived to be for the good of the workers and the organization. Therefore, such actions are perceived to be less unfair and less likely to harm loyalty. The second factor is neutrality, that is, the belief that behavior is unbiased. In conflict resolution among workers, supervisors' active involvement and interference are expected and appreciated in high-power distance and paternalistic cultures, whereas it is perceived to be breaching

supervisors' neutrality in low-power distance and individualist cultures. Finally, the third factor is status recognition, that is, the extent to which decision makers show respect toward employees. People feel valued when they are shown respect, but what constitutes respecting behavior is culturally determined. There is a dearth of cross-cultural research in this exciting area.

See Also the Following Articles

Attitudes ■ Industrial/Organizational Psychology, Overview ■ Organizational Justice ■ Personnel Selection ■ Recruitment ■ Work Motivation ■ Work Teams

Further Reading

Adams, J. S. (1965). Inequity in social exchange. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 267–299). New York: Academic Press.

Allen, N. J., & Meyer, J. P. (1990). The measurement and antecedents of affective, continuance, and normative commitment to the organization. *Journal of Occupational Psychology*, *63*, 1–8.

Blum, M. L., & Naylor, J. C. (1968). *Industrial psychology: Its theoretical and social foundations*. New York: Harper & Row.

Erez, M., & Earley, P. C. (1993). *Culture, self-identity, and work*. New York: Oxford University Press.

Hackman, J. R., & Oldham, G. R. (1980). *Work redesign*. Reading, MA: Addison–Wesley.

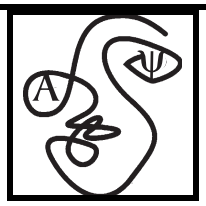
Industrial Democracy in Europe International Research Team. (1981). *Industrial democracy in Europe*. New York: Oxford University Press.

Leung, K. (1997). Negotiation and reward allocations across cultures. In P. C. Earley, & M. Erez (Eds.), *New perspectives on international industrial/organizational psychology* (pp. 640–675). San Francisco: New Lexington Press.

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting and task performance*. Englewood Cliffs, NJ: Prentice Hall.

Organ, D. W. (1988). *Organizational citizenship behavior: The good soldier syndrome*. Lexington, MA: Lexington Books.

Vroom, V. H. (1964). *Work and motivation*. New York: John Wiley.



Industrial/Organizational Psychology, Overview

Sheldon Zedeck

University of California, Berkeley, California, USA

1. Definition
2. History
3. Contents of the Section
4. Other Topics
5. Future Trends
6. Conclusion
   Further Reading

The section on industrial/organizational (I/O) psychology presents a broad overview of the multiple topics studied under its rubric. Some of the topics are quite broad, such as motivation taxonomies, while some are rather specific, such as system safety. The goals of the section are to provide the reader with an introduction to the many areas of concern to I/O psychologists with respect to the areas in which they conduct research and practice. The purpose of this overview article is to (1) define and describe industrial and organizational psychology as a field, (2) provide a brief history of the field, (3) summarize the contents of the section, (4) provide brief summaries of topics not covered in this section, and (5) present one view of the trends in research and practice.

1. DEFINITION

Broadly speaking, psychology is the study of behavior. With this general definition as the basis, industrial/organizational (I/O) psychology can be defined as the study of behavior that takes place within the context of an organization (e.g., a work setting). Adopting such a broad definition provides the opportunity for I/O psychologists to address research and practice questions that most other psychologists study most often in isolation of a context. For example, cognitive psychologists study information processing. An I/O psychologist may also study information processing, but in the context of decision making. Or, a personality psychologist may study the basic dimensions of personality. An I/O psychologist may study personality psychology to determine the traits and characteristics of individuals that will lead to effectiveness as leaders or success in organizations. Yet another example can be drawn from social psychologists, who study small group behavior and attitudes. I/O psychologists could study similar topics, such as team behavior and job satisfaction, respectively. One additional example can be taken from biological psychology, which can focus on stress and circadian rhythms. Here, I/O psychologists might be interested in stress due to conflict between work and family, and in the attitudinal and behavioral outcomes that result from shift work in organizations. These examples demonstrate the broad range of topics that I/O psychologists are concerned with and the issues that they address. They also illustrate the approach taken by I/O psychologists, which is to conduct research that derives principles

©2004 Elsevier Inc.
All rights reserved.

of individual, group, and organizational behavior and to apply that research base to the solution of problems in organizational settings.

One cautionary note about the preceding illustrations: they are not intended to suggest that I/O psychologists predominantly or exclusively address practical, field, or applied issues. Though many I/O psychologists conduct their research in viable organizations, and may do so to address particular problems inherent to those environments, their research is both basic and applied. That is, much of the research that is conducted by I/O psychologists contributes to the understanding of behavior as well as serves to remedy or prevent problems in organizations.

For a more complete overview of the profession of I/O psychology, see the Web site of the professional association for I/O psychologists, the Society for Industrial and Organizational Psychology (www.siop.org). This site provides complete information about the profession, which has over 6,000 members from 50 U.S. states and approximately 40 countries. I/O psychologists conduct research and consult with organizations on the following topics:

Employees:

- Testing
- Selection and placement
- Training and development
- Employee attitudes
- Employee motivation
- Leadership

Organizations:

- Change management
- Organizational climate
- Organizational culture
- Job design
- Job evaluation
- Organizational structure
- Team building
- Decision making
- Workplace planning

Before proceeding with the history of I/O psychology, another term, organization, must be defined. An organization exists when two or more people get together for the purpose of accomplishing a common goal. This is a very broad and general definition of an organization, adopted to highlight the flexibility available to I/O psychologists for their context of study. The organization can be one that is global, national, or local; it can be an organization that is intended to last for a long period of time, or it can be one that is assembled for a short-term mission, such as a political campaign. The members of the organization may be paid for their services or may be volunteers; thus, the organization can be either for profit or non-profit. The members of the organization may be selected or elected to be part of the organization. The organization can be in business for the purposes of being in private industry, the public sector (government agency), or an educational institution.

One final topic for this section of the overview pertains to who is an I/O psychologist. Generally, I/O psychologists receive their training from doctoral level psychology departments that have special programs in I/O psychology. Thus, these individuals generally have a broad background in the research and methodology of psychology as well as in data analysis. Another route for education is to obtain a doctorate in organizational behavior (OB) from schools of business administration. Such programs often interact closely with psychology departments to provide the basic background in psychology needed to conduct research and practice in the domain of behavior. In addition, these programs provide a greater emphasis on business-related topics such as strategic planning, finance, and marketing. Often, however, the perspectives of the I/O psychologist and OB student may vary. For example, the former may be interested in individual differences and micro-level issues regarding how the individual behaves within the context, whereas the latter may prefer to focus on more macro-level issues such as the organization's impact on the individual. Note that this distinction is fuzzy and not rigid—the ideal prototype I/O psychologist is one who understands both micro- and macro-level issues.

2. HISTORY

This section identifies the primary individuals and events that have defined and distinguished the field of I/O psychology. This history is presented in a chronological order to show how the field has developed over the last 100 years and how the field has been influenced by historical and social events.

The formal history of the field begins in the early 1900s. One of the earliest psychologists recognized for his application of psychological principles to problems of the work setting is Hugo Munsterberg. Munsterberg studied in the laboratory of Wilhelm Wundt, an experimental psychologist who studied perception in Germany. When Munsterberg arrived in the United

States, he established a psychological laboratory at Harvard University, where he studied issues concerned with industrial accidents and the relationship between abilities and performance and industrial efficiency. He wrote one of the earliest books on I/O psychology in 1913, *Psychology and Industrial Efficiency*. Perhaps one of his earliest contributions was his applied work, his studies on the characteristics of successful motormen for the Boston streetcar system.

Another early pioneer, who also studied with Wundt, was James McKeen Cattell. Cattell studied individual differences in people; the critical notion for him was that differences between people should not be considered as "errors," which was the way experimental psychologists viewed them, but rather as reliable characteristics that differentiated people and that explained why people differed from each other. Cattell was one of the first to use the term "mental test," and, in many ways, is the founder of the field of selection and personnel.

Two other psychologists who were active in the early 1900s are Walter Dill Scott and Walter Van Dyke Bingham. Scott was active in two areas: he explored the field of advertising as well as the issue of selecting sales personnel. Regarding the latter, he was one of the earliest to demonstrate that the typical interview lacked reliability and validity. He was also interested in training and learning curves, performance rating systems, trade tests, occupational families, and job specifications. Another distinction for Scott is that he was a professor of applied psychology at Carnegie Institute of Technology in the mid-1910s.

Another landmark in the 1910s was the work of Frank and Lillian Gilbreth. They applied the "scientific management" principles of Frederick W. Taylor (1911) to educational institutions. Scientific management basically promoted the simplification of work processes. Work methods were designed to be efficient, and workers were selected who could be trained to perform the tasks. The Gilbreths applied time and motion studies to the ways in which workers performed tasks, both to understand how those tasks were performed and to design a more efficient way in which to perform those tasks in order to reduce fatigue and to increase productivity. These events were the forerunners for work on job analysis as well as human factors engineering—both topics of relevance today for I/O psychologists.

A significant historical event impacted the role of psychologists in the early 1900s—the entry of the United States into World War I in 1917. Scott, along with Bingham, adapted an intelligence test that was designed for individual testing to a format that could be used with groups of test takers; such a test was needed for use with large numbers of candidates to determine who was fit to serve in the armed forces. The result was the Army Alpha test, which was used to select recruits into the armed forces. This was an instance of applying principles of individual differences to selection problems; it was also the forerunner of major efforts to develop selection tests and systems for organizations, a trend that continues today.

The 1930s introduced another trend in I/O psychology: the study of human relations and attitudes. Elton Mayo was interested in studying the relationship between work—particularly boring work—and attitudes and performance. Mayo, Roethlisberger, and Dickson are the main researchers identified with a classic set of studies—the Hawthorne studies—that took place in the Hawthorne plant of the Western Electric Corporation. The study began in 1926 and lasted for more than 10 years. The initial goal of the study was to explore the relationship between productivity and plant conditions, such as lighting, heating, work hours, rest breaks, and other manipulations that were under the control of the organization. The results were contrary to what was expected: even when the conditions were made worse, productivity increased. This finding led the researchers to interview the workers, who indicated that they were "pleased" that the organization was paying attention to them, as represented by the efforts to study the problems; as a result, they increased their performance. This finding—that performance can be increased when workers are attended to—is now known as the Hawthorne Effect. This research led to a trend toward greater concern for the individual workers and an emphasis on social relations, known as the human relations movement. Topics that resulted from this movement include motivation, job satisfaction, and morale.

World War II (early 1940s) was also instrumental in inspiring psychologists to begin new explorations for the purpose of helping the war effort. Of particular significance is the role psychologists played in the design of cockpits for the new fighter planes that were being used in the war. The different types of planes and their cockpit controls required pilots who could function in environments with all sorts of new airplane controls, gauges, and displays. Consequently, the psychologists helped in the design of the cockpit, to facilitate quicker recognition of the dials and controls and with less error in their use. This is another example of human factors engineering as well as work design.

In addition to human factors engineering, World War II was the catalyst for research on selection and assessment. New selection tests were developed to help match recruits to jobs that fit their abilities. In addition, attempts to identify personnel who could serve as spies for the United States resulted in the development of the "assessment center" procedure. The assessment center is a mixture of multiple types of assessment, e.g., individual tests, personality assessment, interviews, role-plays, and group exercises, many of which are simulations of actual job behavior. The set of exercises is designed to provide a picture of the individual and how he or she is likely to behave in a work setting. Today, assessment centers are a major industry used to select people for all levels within an organization.

After WWII, I/O psychology became more recognized, and as a result, degree programs were established at major universities and colleges in the United States. Many of the early pioneers in the field brought their training in experimental or social psychology to the programs, producing new Ph.D.s with interests in applied psychological problems. Some of the issues studied were leadership, participative decision making, and job satisfaction.

The role of the I/O psychologist was again challenged in the 1960s, with the passage of the 1964 Civil Rights Act. This act made it illegal to discriminate in selection, hiring, and promotion of employees based on race, ethnicity, national origin, and gender. As a result of this act, the hiring and promotion practices of many organizations were challenged, which led to the organizations' attempts to demonstrate the validity and usefulness of their selection procedures. For I/O psychologists, this meant increased efforts and research in understanding issues of selection and validation as well as better understanding of how tests can be used in the workplace environment. This law spurred research on validation models, the development of job analysis procedures, concerns about the measurement of performance, and the concept of validity generalization.

In the latter part of the 20th century (1970s to 2000), even though there was still a major focus on personnel psychology, a considerable amount of research and practice focused on the individual within the organization. Topics such as goal setting, satisfaction, organizational commitment, motivation, organizational climate, person–organization fit, organizational citizenship behavior, and justice received increasing research attention. From an economic perspective, some issues of interest were mergers, acquisitions, and downsizing.

The late 1990s and the early part of the 21st century have continued to see the interest in the social aspects of behavior. But this also has been influenced by the changing nature of work and the reliance on technology for that work. This trend will be revisited at the end of this overview. For now, and as a summary, the presentation of this brief history was intended to show that practical problems, within a context, often influenced by politics, economics, conflicts (war), and legislation, have been instrumental in defining what I/O psychologists studied.

3. CONTENTS OF THE SECTION

One way of structuring how I/O psychologists study their topics is to view the study of organizations in three levels or phases along with their associated issues: (1) the need to develop the organization, (2) the need to obtain members for the organization, and (3) the need to maintain (and retain) the organization as a viable entity with members who contribute to its mission. Developing the organization refers to the issues that must be considered when starting an organization. These topics include organizational philosophy, organizational theory, organizational structure, leadership, communications, decision-making strategies, and other issues that relate to how the organization and its jobs will be designed, established, and conducted.

3.1. Development Phase

This section of the encyclopedia contains articles relevant to the development phase on the topics of organizational structure (authored by Tetrick and Camburn), power, authority, and leadership (Munduate Jaca and Medina), organizational culture and climate (Peterson and Fischer), decision making (Jungermann), and organization development (Burke). These are more global topics that focus on organizational characteristics. But there are also very specific concerns, such as how the jobs will be designed and analyzed. A particular article within this section that is relevant to this concern focuses on job analysis, design, and evaluation (Smither). In addition to interest in the tasks needed to be performed and the knowledge, skills, and abilities (KSAs) needed to perform those tasks, organizations are also concerned with the safety of work, which is discussed in articles on system safety (Wilpert) and work safety (Zohar).

The topics described above are critical for the establishment of an organization, since they impact the type

of recruits attracted to the organization and the types of systems that will be developed to maintain the organization. Note, however, that as stated earlier, these are not isolated topics that are restricted to the development phase of setting up the organization. Here, as in the other two phases, the individual topics are appropriate for discussion and consideration for all phases. For example, although it is critical to identify the organizational structure at the outset, concerns about performance and satisfaction could lead to the need for organizational development activities that would take place several years after the establishment of the organization. After studying an organization that has been in operation for several years, it may be necessary to make changes in the structure to accommodate changes in the nature of the workforce, the demands of the public regarding the product produced by the organization, and new laws that govern how products should be produced and what materials they contain. The critical points to appreciate are that all of the topics are interrelated and all are important at any and all phases in the life of an organization.

3.2. Selection Phase

Once an organization is designed and there are plans for its operations, the next issue is the second phase: obtaining members to participate in that organization. When considering how to select employees at all levels within the organization, the I/O psychologist is concerned with the bridge between design of jobs and how recruits are selected to perform those jobs; that bridge is based on the previously mentioned article on job design and analysis. In addition, there is the related issue of how employees will be compensated for the work they perform in their roles, which is discussed in an article on compensation (Thierry).

Once jobs are designed, it is necessary to understand how employees will be evaluated in their positions or jobs—the topic of performance appraisal. Here, I/O psychologists need to understand what contributes to success and effectiveness in the job and organization, and how to measure employees on the relevant criteria. Two articles deal with the measurement of individual and organizational performance on the tasks: one on competence at work (DeNisi) and one on productivity (Pritchard). High individual performance and high organizational productivity are the criteria that the organization selects employees to achieve. Another criterion used to measure the organization's success is turnover; this is discussed in the article on labor turnover (Hom). The selection system is designed to identify those who can meet these goals.

This section also contains three critical articles that set the stage for the other topics within the selection phase—occupational choice (Hansen), recruitment (Breaugh), and personnel psychology (Chan). The occupational choice article considers how people choose the careers they want to pursue, focusing on the individual and what he or she wants to do as a career. The personnel psychology article focuses on the organization's choice of individuals, which is obviously related to what individuals decide to do in terms of careers. How the organization goes about identifying candidates for its positions is discussed in the article on recruitment. In other words, there is a reciprocal relationship between one's choice of a career and the organization's recruitment and choice of people.

Once the organization is ready to implement a selection system, it needs to consider the issues discussed in the articles on employment discrimination (Tenopyr) and employment interviewing (McDaniel and Whetzel). The former deals with laws that exist, particularly in the United States, that influence what selection devices can be used and under what conditions, as well as the litigation surrounding the use and validation of the tests. Related to the topic of discrimination and selection is the issue of affirmative action, which is discussed in an article by Kravitz. The topic of validation is covered in the previously mentioned article on personnel psychology (Chan) as well as in the article entitled "Personnel Selection" by Ones, which appears in another section of the encyclopedia.

A link between obtaining members for and maintaining the organization is the topic of training and development. Organizations can select employees who have the aptitude to be trained, or employees who are already trained and capable of producing and performing rather quickly after entry into the job. Training is a critical issue of concern for I/O psychologists. A special topic related to training is that of executive coaching and executive development, discussed in an article by Silzer, which is concerned with working with employees to enhance and develop their skills for management positions. This not only is an issue when deciding who to select and whether training is needed, but also can serve as a motivational tool for employees.

3.3. Maintenance Phase

Once the organization has determined how it will obtain its employees and implements the systems needed

to obtain those members, it turns its attention to the maintenance phase: how to retain the members in the organization and to further maintain and develop the organization to sustain itself in the marketplace. This phase is concerned with issues such as reward systems, motivation, job satisfaction, organizational climate and culture, work and family, conflict, and stress.

Articles relevant to the maintain phase are those on the topics of motivational taxonomies (Vancouver), motives and goals (Kleinbeck), and boredom (Balzer, Smith, and Burnfield). These articles address issues such as how to motivate employees to perform better and how to make the work more motivating and satisfying for the employees. Articles that deal with group and team issues include those on organizational participation (Heller), organizational socialization (Bauer), and cooperation at work (Salas). These issues relate, in part, to whether there is a person–environment fit, discussed in an article by Ostroff and Aumann that deals with selection, motivation, organizational characteristics, recruitment, and other factors that determine whether the organization's choice and the individual's choice are consistent. There may be problems, or conflict, discussed in an article by Castro and Ryan.

While attempting to maintain and further enhance the organization's position in the marketplace, the organization can undertake programs designed to assess how well it is functioning. This issue is discussed in the article on organizational diagnosis (Bussing). Another article in this section pertains to downsizing and outplacement (Cascio), a topic important and timely in today's economic environment. Organizations are merging and downsizing in order to achieve greater economies of workforce and productivity.

Note that the above trichotimization into development, selection, and maintenance phases is not exhaustive or mutually exclusive; rather, the topics are integrated. That is, it would be difficult to consider issues of selection without considering issues of training and motivation. This means that if an organization determines that it needs to place employees into position "X," it could do so by selecting those who already have the "X" abilities and skills; it also means their compensation would be affected by the fact that the hired are skilled. In contrast, the organization could decide to hire those individuals who demonstrate the aptitude or ability to learn and gain the abilities and skills needed to perform the tasks. This would mean that there would be greater emphasis on training those hired, which in turn might mean that their compensation could initially be lower because they are relatively less skilled. For both types of hires, however, motivation is critical. How do organizations motivate workers to be fully productive? Does the fact that the organization provides the training to relatively less skilled employees mean that those employees will be more committed to the organization, and consequently the organization will have a lower turnover rate? These issues and questions need to be considered jointly. Although the articles in this section treat the topics separately, the challenge for the reader is to recognize the necessary integration.

4. OTHER TOPICS

An encyclopedia often provides information on all topical phrases that are used within the discipline. Those who are somewhat familiar with I/O psychology might wonder why there are no specific articles in this section on topics such as job satisfaction, incentives, test fairness, and work performance. The reason for the omission of articles on such specific topics is because their essence is conveyed in other articles appearing in this or other sections of the encyclopedia. For the topics just mentioned, for example, the reader is referred to the articles on boredom, motivational taxonomies, personnel psychology, and competence at work, respectively. The point is that this section on I/O psychology is reasonably complete and provides an overview of the major topics and issues in the field.

5. FUTURE TRENDS

An obvious point of interest for readers of this section is what the future holds for the field of I/O psychology. Our crystal ball suggests that the major topics covered in this section will remain the same, but that the nature of work—how it is performed, what it is composed of, and where it is performed—will change. That is, I/O psychologists will always be interested in how organizations should be developed, how to obtain people for those organizations, and how to maintain and promote the survival and growth of the organization. But what people do in the form of work, and how and where they perform that work is likely to change.

To elaborate, let us return to the history of the field. Reviewing the "older" textbooks, such as one by Viteles, shows that their tables of contents cover many of the topics included in this section of the encyclopedia. Among the 27 chapters of the Viteles text are chapters

on individual differences, the social foundations of industrial psychology, vocational selection, job analysis, the interview, tests for different skills and industries, safety, training methods, industrial fatigue, motives, monotony, and supervision. The topics have remained the system—the way that I/O psychologists have approached the research problems and analyzed the data have become more sophisticated and elegant.

So, what has changed and what will continue to change? The answer is that (1) work, (2) the worker, and (3) the context will change. First, the nature of work will change. It is obvious that more and more work will be driven by computer technology and that employees will be producing less manually. More and more employee work will focus on providing information and services. Thus, I/O psychologists need to be concerned with systems that can describe the tasks to be performed (job analysis and job design), identify the KSAs needed to perform those tasks (personnel psychology), and figure out how to measure effectiveness on those tasks (competence at work). Work will become more fluid; jobs will not be well-defined. Much of this work is likely to be done in groups and teams, which means emphasis on how to select and motivate team members. Also, it is likely that much of the work will be done in virtual teams. An obvious implication of all of this change is that the relationship between the worker and the client will change. There will be less direct contact and more reliance on the worker fulfilling the client's expectations without being on site to push for services.

Second, workers may interact differently with their peers and supervisors. More and more meetings may take place via teleconferencing instead of face-to-face. Accordingly, I/O psychologists will need to pay more attention to cross-cultural psychology, to topics of personality and interpersonal relations, and to work and non-work balance.

The third issue that will impact I/O psychology, context, is concerned with the globalization of doing business and multiculturalism of the workforce. I/O psychologists will need to address the performance of work in the context of the globalization of the economy with a workforce that is more culturally diverse than it was in the past. This means that workers will need to be more adaptable to where they work and with whom they work. Not only will workers be asked to interact with others from different cultures and backgrounds, but they also may be asked to work in locations foreign to them (e.g., other countries).

The challenge for the I/O psychologist will be to develop systems (selection, training, compensation, motivation, etc.) that recognize the changes in the work that is being performed, how it is being performed, who is performing it, and in what context. Workers will need to be more flexible and agile, and work at faster paces and speeds. Workers may need to be more self-driven and work in contexts in which they are empowered and in which there will be self-learning. Workers will need to adjust to uncertainty and the temporary nature of their work. They will need to adapt to flexible work patterns. And, they may need to assume that they will have multiple careers over their life.

6. CONCLUSION

This section focuses on the study of behavior for those who operate within an organizational context. It presents the issues and interests of I/O psychologists as they attempt to study how work is performed, who performs it, and where it should be performed. The topics contained within this section are integrated and should be studied in relation to each other—focusing on one article without considering others will lead to inappropriate assessments, conclusions, and solutions. Hopefully, the section conveys the message that the study of work and workers is complex and that many variables need to be considered when attempting to provide solutions to work problems. The literature shows that less than 50% of the variability in individual and group performance can be explained by the variables I/O psychologists study as they attempt to explain and predict that behavior. The goal for the future is to explain more of this variability.

See Also the Following Articles

Decision Making ■ Employment Interviewing ■ Industrial/Organizational Psychology across Cultures ■ Organizational Culture and Climate ■ Organization Development ■ Organizational Socialization ■ Personnel Selection ■ Power, Authority, and Leadership ■ Recruitment ■ Work Environments ■ Work Motivation ■ Work Role, Values Sought in the ■ Work Teams

Further Reading

Benjamin, L. (1997). A history of Division 14 (The Society for Industrial and Organizational Psychology). In D. A. Dewsbury (Ed.), *Unification through division: Histories of the divisions of the American Psychological Association*

(Vol. 2, pp. 459–466). Washington, D.C: American Psychological Association.

Goldstein, I. R., & Ford, J. K. (2002). *Training in organizations* (4th ed.). Belmont, CA: Wadsworth.

Howard, A. (Ed.) (1995). *The changing nature of work*. San Francisco, CA: Jossey-Bass.

Ilgen, D. R., & Pulakos, E. D. (Eds.) (1999). *The changing nature of performance*. San Francisco, CA: Jossey-Bass.

Katzell, R. A., & Austin, J. T. (1992). From then to now: The development of industrial-organizational psychology in the United States. *Journal of Applied Psychology*, *77*, 803–835.

Koppes, L. L. (1997). American female pioneers of industrial and organizational psychology during the early years. *Journal of Applied Psychology*, *82*, 500–515.

Landy, F. J. (Ed.) (1993). Early influences on the development of industrial and organizational psychology. In, *Exploring applied psychology* (pp. 81–118). Washington, D.C: American Psychological Association.

Landy, F. J. (1997). Early influences on the development of industrial and organizational psychology. *Journal of Applied Psychology*, *82*, 467–477.

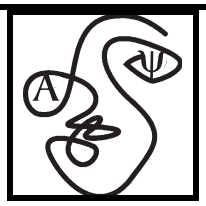
Landy, F. J., & Conte, J. (2004). *Work in the 21st Century: An introduction to industrial and organizational psychology*. New York: McGraw-Hill.

Roethlisberger, F. J., & Dickson, W. J. (1939). *Management and the worker*. Cambridge, MA: Harvard University Press.

Munsterberg, H. (1913). *Psychology and industrial efficiency*. New York: Houghton Mifflin.

Taylor, F. W. (1911). *Principles of scientific management*. New York: Harper.

Viteles, M. S. (1932). *Industrial psychology*. New York: W. W. Norton.



Injury in Sport

Albert J. Petitpas and Britton W. Brewer
Springfield College, Springfield, Massachusetts, USA

1. Introduction
2. Psychosocial Risk Factors for Sport Injury
3. Psychological Effects of Sport Injury
4. Sport Injury Prevention
5. Psychological Interventions in Sport Injury Rehabilitation
Further Reading

GLOSSARY

athletic identity The degree to which one identifies with the athlete role.
proprioceptive Sensory or perceptual awareness of body movement and position, particularly with respect to the limbs.
psychosocial risk factors Physical and psychological conditions existing prior to injury that predict the occurrence of injury.
rehabilitation adherence The extent to which one follows a prescribed rehabilitation protocol.
sport injury Physical injury sustained during participation in recreational or competitive sport activity.
tough-mindedness A personality orientation characterized by assertiveness, self-assurance, independence, and a tendency to place stress on oneself.

Psychological factors are important elements in most aspects of the injury process in athletes. Psychological factors have been identified in both the occurrence and aftermath of sport injury, and psychological interventions have been used to reduce injury risk and to facilitate physical and psychological recovery from sport injury.

1. INTRODUCTION

Physical injury is a common by-product of sport participation. Literally millions of athletes around the world sustain one or more injuries each year, incurring billions of dollars of medical costs in the process. Although most of the injuries are minor and require minimal medical treatment, many of the injuries require extensive and costly diagnostic testing, surgery, and rehabilitation. In the short term, sport injuries can impair performance both on and off the athletic field. Athletes can be prevented from training or competing as a result of injury, and their daily activities can be disrupted by injury-related pain or disability. In the long term, certain sport injuries may place athletes at increased risk for developing arthritis and other chronic medical conditions. Some of the more severe injuries may bring about the end of athletes' sport careers and result in permanent disabilities.

Historically, injury has been considered primarily as a physical phenomenon within the sport community, with an emphasis on identifying the physical factors that cause injuries to occur and on helping athletes to recover from injuries. Unfortunately, even with all of the medical and technical advances that have been made over the past several decades in equipment, playing surfaces, and other safety features, the number of sport-related

©2004 Elsevier Inc.
All rights reserved.

injuries has continued to increase. This unexplained rise in injury rates has been accompanied by a growing recognition of the possible role of psychological factors in the sport injury process. Psychological factors are now thought to contribute to not only the occurrence but also the rehabilitation of sport injuries. Indeed, psychological interventions have been implemented to assist in the prevention and treatment of sport injury.

2. PSYCHOSOCIAL RISK FACTORS FOR SPORT INJURY

Athletes and coaches are well aware of physical factors that can precipitate injury, including overuse, muscular tension, and trauma. Psychosocial factors can also place athletes at increased risk for injury. Over the past three decades or so, elevated life stress has emerged as a highly consistent predictor of sport injury. In American football and a variety of other sports, athletes who have recent histories of stressful life events are more vulnerable to injury than are those without such stressful experiences. Life events perceived as negative (e.g., loss of a loved one, getting fired from a job) typically have the most adverse impact on the occurrence of sport injury, although even events perceived as positive (e.g., academic achievements, new romantic relationship) can be stressful and can result in increased injury risk.

In addition to life stress, personality and coping resources are psychosocial factors that can affect susceptibility to sport injury. Injury risk appears to be elevated for athletes low in social support or high in trait anxiety, Type A behavior, negative mood, and/or tough-mindedness. Athletes who are not satisfied with the amount of technical assistance, informational guidance, and emotional comfort they receive from people important to them are at increased risk for injury. Similarly, athletes who are generally anxious, hard driving, hurried, hostile, assertive, self-assured, depressed, and/or angry are more vulnerable to injury than are those without such characteristics. Personality and coping resources can also influence the impact of life stress on the occurrence of sport injury. For example, the adverse effect of stressful life events on sport injury is especially pronounced for athletes who tend to be anxious in competitive sport situations, lack social support, and/or have difficulty coping with stress.

It is not known exactly how life stress, personality, and coping resources affect vulnerability to sport injury, although there is reason to believe that both cognitive and physiological factors are directly involved in injury occurrence. Life events, both positive and negative, often elicit stress reactions that can cause athletes' muscles to be tense, peripheral vision to narrow, and attentional focus to diminish, all of which may place athletes at increased risk for injury during sport participation.

3. PSYCHOLOGICAL EFFECTS OF SPORT INJURY

In addition to the obvious detrimental effects of sport injury on physical functioning, sport injury can have a deleterious effect on various aspects of psychological functioning. In particular, injury can have an adverse impact on the thoughts, feelings, and actions of athletes. With respect to thinking, sport injury is associated with reduced self-esteem and physical self-confidence as well as with elevated confusion. From an emotional standpoint, it is not uncommon for athletes to experience negative feelings such as anger, confusion, depression, fear, and frustration following injuries. Although clinically meaningful levels of emotional distress have been found in approximately 5 to 24% of athletes with injuries, most negative emotions tend to dissipate over the first month following the injuries. Once athletes adjust to the initial changes caused by their injuries, they are typically able to focus primary attention on their rehabilitation programs. For example, injury appears to spur athletes' use of active coping strategies such as vigorously pursuing rehabilitation goals and seeking support or information from others. As athletes become more invested in their rehabilitation regimens, their moods return to typical preinjury levels.

Athletes' psychological responses to injuries are influenced by characteristics of the athletes themselves and the situations in which they sustain their injuries, including characteristics of the injuries themselves. Athlete characteristics associated with emotional reactions to injury include age, athletic identity, and hardiness, with athletes who are younger, more strongly identified with the athlete role, and less hardy reporting greater emotional disturbance following injuries. The extent to which athletes adhere to their injury rehabilitation programs—a behavioral response to injury—is related to athlete characteristics such as internal health locus of control, pain tolerance, self-motivation, and tough-mindedness. Athletes who consider themselves

responsible for maintenance of their health, tolerate pain well, and are self-motivated and tough-minded tend to demonstrate high levels of adherence. Injury is also a prominent factor leading to dropout in youth sport participation.

With respect to characteristics of the situations in which injuries occur, athletes report greater postinjury emotional disturbance when they perceive their injuries as severe, their life stress as high, their rehabilitation progress as poor, and their social support as low. Athletes tend to stick with their prescribed rehabilitation programs better when they perceive their injuries as severe, their level of social support as high, the treatment as effective, the clinical environment as comfortable, and the rehabilitation appointments as conveniently scheduled. Emotional and behavioral responses to injury are also related to each other; that is, athletes with high levels of mood disturbance tend to adhere well to their rehabilitation programs, whereas athletes with low levels of mood disturbance tend to adhere poorly to their rehabilitation programs.

Athletes who incur sport career-ending injuries are vulnerable to considerable psychological distress, particularly if they are competing at a high level and view themselves primarily or exclusively as athletes. Intercollegiate athletes who sustain career-ending injuries tend to be less satisfied with life 5 years after the conclusion of their playing careers, particularly if they believe that they have not achieved their sport goals.

Individuals incurring catastrophic sport injuries that result in permanent functional disabilities, typically from damage to the head or spinal cord, must make profound physical, emotional, social, and lifestyle adjustments. Although catastrophic injuries are rare occurrences, they happen with enough frequency that the National Center for Catastrophic Sports Injury Research was established in 1982 to track prevalence rates and make recommendations for safety and prevention that have been implemented in most contact sports as well as in gymnastics and cheerleading.

In addition to injury severity, other aspects of the injury itself—onset, course, history, and type—can affect psychological responses to sport injury. The onset of a sport injury can be either acute or gradual. Sudden and unexpected injuries are classified as acute and often place different demands on athletes' personal coping resources and elicit a different response from their social support systems compared with nagging or overuse injuries that eventually cause athletes to stop competing due to the gradual accumulation of physical damage or discomfort.

Sport injuries can also be classified by the course to recovery. Progressive-type injuries show a continual stepwise process toward either full recovery or permanent disabilities. The speed of the progression may vary, but the course is predictable with few surprises or unknowns. Constant-type injuries require athletes to adjust to physical impairments that remain stable or invariant over time. In episodic-type injuries, athletes must learn to cope with injuries that recur or flare up periodically and unexpectedly. Injury course is important because individuals' perceptions of control and predictability in various situations have been linked to the amount of anxiety that they manifest in those situations.

History and type of injury can also affect athletes' psychological responses to injury. Athletes who have previous histories with the same injury tend to be more knowledgeable of what to expect and may have less anxiety than do athletes who experience the injury for the first time. Athletes with injuries of a type that requires a cast or involves some form of visible impairment may experience different reactions from their social support networks compared with athletes with nonvisible injuries who are forced to answer repeated questions about why they are not competing.

The ways in which athletes respond psychologically to injury may have important implications not only for the athletes' quality of life following injury but also for their physical recovery. As shown in Table I, a variety of psychological factors are related to sport injury rehabilitation outcomes. In general, athletes who recover faster or more effectively tend to be those who demonstrate a favorable mental health profile, receive adequate social support, and adopt an approach to their rehabilitation in which they maintain a positive attitude, apply psychological coping strategies, take responsibility for achieving desired rehabilitation outcomes, and adhere to their rehabilitation programs.

Athletes' readiness to resume training postinjury requires both physical and psychological assessment. Warning signs that could help to identify athletes who are not ready psychologically to return to competition include dwelling on minor somatic complaints just prior to the recovery target date, withdrawing suddenly from teammates and/or coaches, exhibiting striking changes in affect or behavior, and engaging in exaggerated storytelling or bragging about accomplishments in and out of sports.

TABLE I
Examples of Psychological Factors Associated with Sport Injury Rehabilitation Outcomes

Variable type	*Variable*
Athlete factors	Age
	Athletic identity
	Gender
	Optimism
	Personality
Situational factors	Social support
Cognitive factors	Attitude toward rehabilitation
	Attributions for recovery
	Goal-setting use
	Imagery use
	Injury appraisals
	Pain coping
Emotional factors	Mood disturbance
	Psychological distress
	Rehabilitation anxiety
Behavioral factors	Adherence to rehabilitation
	Physical activity
	Social support seeking

4. SPORT INJURY PREVENTION

Just as psychological factors can contribute to sport injury, psychological treatments can help to prevent sport injury. Initial evidence of the potential preventive impact of psychological interventions on the occurrence of sport injury was obtained almost by accident as an unintended side effect of interventions designed to enhance sport performance. Training athletes in attentional control, biofeedback, imagery, and relaxation can reduce conditions (e.g., distraction, muscular tension, perceived stress) that not only hamper sport performance but also increase vulnerability to injury. Teaching athletes psychological strategies for managing stress (due to both positive and negative life events) can lower their risk of sustaining sport injuries. In situations where athletes have experienced successful competitive performance, such interventions can help athletes to reduce their injury risk by maintaining their focus and avoiding complacency.

Developing athlete social support groups that provide forums for individuals to share their life events, both positive and negative, may buffer the amount of stress induced by these experiences and thus moderate injury susceptibility. Peer mentoring systems within teams or athletic departments can also strengthen the overall social support system and offset the potentially adverse consequences of stress.

5. PSYCHOLOGICAL INTERVENTIONS IN SPORT INJURY REHABILITATION

As shown in Table II, a variety of psychological interventions have been successfully implemented to alter the psychological responses, rehabilitation processes, and rehabilitation outcomes of athletes with injuries, particularly knee injuries. As with interventions targeted at preventing sport injuries, the reasons why the interventions work are not fully understood. However, most of the interventions are thought to increase motivation for rehabilitation, and this may in part account for the therapeutic effects of the interventions. It has also been proposed that biofeedback enhances athletes' awareness of their body movements and positioning and that relaxation and imagery influence physiological processes central to healing such as immune/inflammatory responses and the regeneration and repair of body tissues.

TABLE II
Psychological Interventions in Sport Injury Rehabilitation

Intervention	*Therapeutic outcome*
Biofeedback	Knee/Leg strength
	Electromyogram (EMG) output
	Range of motion
Goal setting	Rehabilitation adherence
	Rehabilitation self-efficacy
	Knee strength
Relaxation/Imagery	Pain
	Reinjury anxiety
	Knee strength
Positive self-talk	Leg strength
Multimodal treatment package	Physical recovery
	Pain
	Anxiety
	Positive mood
	Positive attitude

See Also the Following Articles

Goal Setting and Achievement Motivation in Sport ■ Overtraining and Burnout in Sports ■ Performance Slumps in Sport: Prevention and Coping

Further Reading

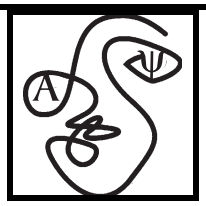
Brewer, B. W. (2001). Psychology of sport injury rehabilitation. In R. N. Singer, H. A. Hausenblas, & C. M. Janelle (Eds.), *Handbook of sport psychology* (2nd ed., pp. 787–809). New York: John Wiley.

Pargman, D. (Ed.). (1999). *Psychological bases of sport injuries* (2nd ed.). Morgantown, WV: Fitness Information Technology.

Petitpas, A., & Danish, S. J. (1995). Caring for injured athletes. In S. M. Murphy (Ed.), *Sport psychology interventions* (pp. 255–281). Champaign, IL: Human Kinetics.

Wiese-Bjornstal, D. M., Smith, A. M., Shaffer, S. M., & Morrey, M. A. (1998). An integrated model of response to sport injury: Psychological and sociological dimensions. *Journal of Applied Sport Psychology*, *10*, 46–69.

Williams, J. M. (2001). Psychology of injury risk and prevention. In R. N. Singer, H. A. Hausenblas, & C. M. Janelle (Eds.), *Handbook of sport psychology* (2nd ed., pp. 766–786). New York: John Wiley.



Intelligence and Culture

Fons J. R. van de Vijver
Tilburg University, Tilburg, The Netherlands

1. Culture and Intelligence: An Emerging Field
2. Three Views on the Relationship between Intelligence and Culture
3. The Measurement of Intelligence in a Cross-Cultural Context
4. Formal Versus Informal Studies of Intelligence
5. Everyday Definitions of Intelligence
6. Conclusion
Further Reading

GLOSSARY

absolutism The viewpoint that holds that cognitive processes are the same across the globe and that emphasizes the context-independent nature of cognitive processes and basic intellectual processes.

bricoleur A metaphor often underlying studies in the informal tradition; a bricoleur (i.e., "jack-of-all-trades") is not interested in a solution that is correct as defined by an external standard (which is typically the case in the formal tradition) but rather is interested in a solution that solves a practical problem (e.g., a leaking roof) using available means.

culture-fair test A test that has been designed in such a way that culture does not differentially influence performance on the test (i.e., item formats are so easy that everyone is supposed to not experience any problem with the test); the concept has been abandoned as unattainable.

culture-free test A test that has been designed in such a way that culture does not influence performance on the test; the concept has been abandoned as unattainable.

everyday intelligence The study of the relationship between culture and intelligence by the observation of intellectual processes in everyday behavior (e.g., arithmetic behavior of shoppers who compare bargains).

formal tradition The study of the relationship between intelligence and culture by the administration of psychometrically adequate tests in different cultures; the scientific approach is seen as the normative model of good problem solving, with an emphasis on the solution of formalized problems that are unlikely to be met in everyday life and on the correctness of solutions.

informal tradition The study of the relationship between intelligence and culture by studying how people solve everyday intellectual problems; the "bricoleur" (i.e., jack-of-all-trades) is seen as the implicit model of problem solving.

nature–nurture debate The scientific controversy about the question as to what extent cross-cultural differences in scores on intelligence tests are due to genetic factors (nature) or environmental factors (nurture).

relativism The viewpoint that holds that intelligence comes into existence in a specific cultural context and that intelligence and culture are so closely linked that any attempt to compare intelligence across cultures is futile.

universalism The viewpoint that basic cognitive processes are universal but that manifestations of these processes may well vary across cultures.

This article discusses three views on the relationship between culture and intelligence. The first, called the relativist viewpoint, holds that intelligence comes into existence in a specific cultural context and that intelligence and culture are so closely linked that any

©2004 Elsevier Inc.
All rights reserved.

attempt to compare intelligence across cultures is futile. The second, called the absolutist viewpoint, maintains that cognitive processes are the same across the globe and emphasizes the context-independent nature of cognitive processes and basic intellectual processes. The third, called the universalist viewpoint, also holds that basic cognitive processes are universal but argues that manifestations of these processes may well vary across cultures. Assessment has always been an important part of the work on intelligence. This article provides a brief history of cross-cultural assessment of intelligence, ending with an overview of what are currently seen as major issues. Then, it presents a description and examples of two kinds of studies of the relationship between culture and intelligence: formal (usually based on a comparison of scores on intelligence tests) and informal (based on the observation of everyday behavior). Then, the article discusses everyday definitions of intelligence and argues that, particularly in non-Western countries, social aspects such as obedience and social skills are seen as part of intelligence alongside reasoning and memory. The article concludes that the universalist position, according to which the basic constituents of intelligence are panhuman, is best supported. However, within predefined limits, cultures define the way in which these constituents are merged into ability patterns.

1. CULTURE AND INTELLIGENCE: AN EMERGING FIELD

It was more than a century ago that the French psychologist Alfred Binet was asked to devise a test that could be used to inform teachers about which pupils have learning difficulties. Tests to predict school performance nowadays are known as intelligence tests and are one of the most successful and widespread applications of psychological assessment. Intelligence test scores were, and still are, the best predictors of success in school and at work. The success has undoubtedly added to the received view that intelligence tests measure intelligence both within and across cultures. However, on closer scrutiny, the relationship between an intelligence quotient (IQ) score and the theoretical concept of intelligence turns out to be complex. At the beginning of the 20th century, Porteus set out to measure the cognitive skills of various cultural groups. He administered his Maze Test, a paper-and-pencil instrument in which the testee must find a way out of a schematically drawn maze, to various cultural groups, including the Bushmen in the Kalahari Desert. Porteus concluded that the Bushmen's visuospatial skills (an important aspect of intelligence) were poorly developed, obviously not realizing that his conclusion clearly contradicted the superb tracking skills of Bushmen in the relatively cueless environment of the Kalahari.

Two further examples are presented to illustrate the close link between intelligence and culture. In the cognitive anthropological literature, there are various examples of expert cognitive performance. For example, the Trukese of the Pulawat atoll in the South Pacific have remarkable navigating skills; they sail over hundreds of miles of open ocean without compasses, sextants, or star tables. The second example comes from studies on Quranic schooling. Some Muslims learn (parts of) the Scripture by heart. In particular, when the learner does not speak Arabic (the language of the Scripture), massive training is needed. This is done by using an incremental learning method in which the first word is learned first, followed by the first and second words, followed by the first, second, and third words, and so forth. Studies have shown that this training leads to improvements of these people in learning in an incremental paradigm. However, their memory skills are not remarkable on tasks that do not require this type of learning. Cultural practice implies frequent exposure to a very specific task, and this leads to the advanced development of a skill.

2. THREE VIEWS ON THE RELATIONSHIP BETWEEN INTELLIGENCE AND CULTURE

With some alleged simplification, one could say that in cross-cultural psychology, three kinds of views can be found on the relationship between intelligence and culture: relativist, absolutist, and universalist.

2.1. Relativist Viewpoint

According to the relativist viewpoint, intellectual processes are closely linked to their cultural context and can be studied only with explicit reference to this context. The visuospatial skills of the Bushmen play an important role in tracking game needed for their everyday survival. The skills of Western individuals are developed to survive in a completely different environment. Therefore, the skills are embedded in different

contexts in both cultures. As a consequence, the skills are not comparable across these two cultures, and it is impossible to design an instrument that can be used to compare these skills across the cultures. Furthermore, there is no need to compare these skills given that all cultural groups are well adjusted to their natural habitats; one cannot say that Bushmen are better (or worse) adapted to the desert than Westerners are to the urban industrialized context. In this view, as socialization progresses, cognitive skill patterns of individuals from different cultures become increasingly dissimilar. Each culture produces its own type of intelligence that is well suited for everyday functioning in that specific cultural context. In this framework, it is important to identify these patterns: How does Bushman intelligence differ from Japanese intelligence?

2.2. Absolutist Viewpoint

The absolutist viewpoint is often associated with a small group of influential researchers such as Arthur Jensen, the late Hans Eysenck, Richard Lynn, and Philip Rushton. It is the opposite of the relativist viewpoint; absolutists argue that cognitive processes are the same across the globe. They emphasize the context-independent nature of cognitive processes. Basic cognitive operations, such as transitivity (e.g., "If John is taller than Mary and Mary is taller than Peter, then John is taller than Peter") and syllogisms (e.g., "If all humans are mortal and Socrates is a human, then Socrates is mortal"), are independent of their context of occurrence. Because of this independence, tests can be designed in such a way that they have a panhuman validity. A test of intelligence that is both reliable and valid in San Francisco will also be valid in the Australian outback and in Sub-Saharan Africa (assuming that the test does not measure information that is available only to a single cultural group). It is clear that with such a strong confidence in the value of intelligence tests, far-reaching conclusions about intellectual skills of various cultural groups can be drawn. Absolutists often assume that these differences are biologically rooted; not surprisingly, they often use the term "race" to describe these differences. Jensen has provided much evidence showing that European Americans show higher IQ scores than do African Americans (approximately 15 IQ points on average). Eysenck has argued that if the races are split up according to skin color, the "yellow race" has the highest IQ scores, followed by the "white race" and then the "black race."

2.3. Universalist Viewpoint

The universalist viewpoint holds that the same basic cognitive structure of humankind is shared by all cultural groups. Cognitive skills and structures, such as reasoning and memory, are present in all cultures, and these skills are called universal. Still, it is clear from cross-cultural studies that culture exerts a powerful influence on intellectual functioning. Culture defines the context of intellectual functioning by determining what, when, and how individuals will learn. A universalist perspective on the navigational skills of the Trukese would hold that these skills consist of universal building blocks that are combined and refined in a lengthy learning process, with the skill being learned by observing experts. In a universal perspective, there are no different types of intelligence. However, cultures differ in terms of which building blocks are used to a greater or lesser extent, and this could lead to cross-cultural differences in mean scores on the more extensively trained skills. Cultures also differ in the specific skills taught (e.g., navigational skills among the Trukese, computer games in Western societies). Such complex skills are based on the universal building blocks that are applied to cultural knowledge.

It is probably fair to say that the universalist viewpoint is the most popular one in cross-cultural psychology. A basic problem of the relativist position is that all cognitive processes are assumed to be context bound. If one is interested in a comparison of cognitive processes across various cultures, the position that these processes are not comparable by definition is not very fruitful. The main problem with this position is the implied overstatement of what could be a valid position in some domains of cognition. Take short-term memory span as an example. A number of items, usually digits, are presented at a rate of one item per second, and the respondent must repeat the items in the same order. The test starts with short series and then continues until the respondent can no longer recall the items correctly. There are strong indications that at least within school-going populations, short-term memory span is relatively invariant across cultures. A relativist viewpoint can quickly become counterproductive when studying these basic universalities in human cognitive functioning. The other extreme position, absolutism, can be equally counterproductive. It is difficult to understand how one could measure visuospatial skills of Bushmen and New Yorkers with one and the same test that gives a valid reflection of their skill levels and that could give a good prediction of real-life behavior in both

environments in which these skills are relevant. In general, it is naive to assume that Western cognitive tests "travel well" and that they easily allow for direct numerical score comparisons of different cultural groups. The popularity of the universalist viewpoint can be easily derived from the shortcomings of the relativist and absolutist viewpoints. The empirical record shows clearly that neither the culture-bound nature of the relativist viewpoint nor the essential denial of cultural influences on cognitive processes is strongly supported. The universalist viewpoint, which is more prudent, has the most face validity.

3. THE MEASUREMENT OF INTELLIGENCE IN A CROSS-CULTURAL CONTEXT

The measurement of intelligence has always played an important role in the literature. If one works in a relativist tradition, one must make sure that his or her assessment of intelligence is adequate for the cultural context. A measure of vocabulary should provide a score that gives an adequate picture of the testee's knowledge of the lexicon. Within the universalist tradition, one must deal with an additional question: Is a test (possibly translated or adapted) adequate to be used in different cultures? It cannot be taken for granted that a test (or its items) developed in one cultural context (typically a Western context) can be simply translated and employed in a different cultural context without any problems. For example, words in a vocabulary subtest might be inadequate or relatively easy or difficult for some cultural groups as compared with other cultural groups. In a Dutch study involving Turkish and Moroccan migrant children and Dutch autochthonous children, it was found that the word "bacon" was much more difficult for the migrant groups, presumably due to the food taboo on pork meat in Islam. The adequacy of a test to measure intelligence in a cross-cultural context cannot be assumed but must be demonstrated.

The question of whether intelligence tests can be applied in different cultural contexts is answered in different ways by relativists, universalists, and absolutists. Relativists argue that basic cognitive processes come into existence in a specific cultural context and that process and context constitute each other; in this view the search for a cross-culturally applicable intelligence test is needless and pointless. Within the absolutist tradition, and particularly within the universalist tradition, the view is more common that there are good and bad intelligence tests, from a cross-cultural perspective, and that it is possible in principle to develop cross-culturally adequate intelligence tests.

Although the terms of relativism, absolutism, and universalism are recent, the theme of how to design intelligence tests that are adequate in various cultures is old. It was recognized a long time ago that familiarity with test stimuli (e.g., figures and words used in a test) could play a role in assessment; that is, people more familiar with the stimuli can be expected to obtain higher scores. Back in the 1940s, Cattell proposed using very simple stimuli in cognitive tests to deal with the problem that various groups were not equally familiar with test material. This led to the development of "culture-free tests." However, the assumed independence of test performance and, among other things, socioeconomic status could not be demonstrated. Subsequently, "culture-fair tests" were developed. Not surprisingly, the existence of culture-free and culture-fair tests has come under critical scrutiny.

Tests that claim to be culture free or culture fair must deal with two related issues. The first is the problem of differential stimulus familiarity. If cognitive tests are administered to cultural groups with widely different educational backgrounds, differences in stimulus familiarity are nearly impossible to avoid and overcome. The second issue is the implicitness of cultural knowledge in tests. Tests and testing situations imply much cultural knowledge, even when this knowledge is not the topic of a test. For example, a situation in which an individual asks questions with answers he or she knows and expects the testee to answer the questions without the help from anyone else is uncommon in everyday life, with experience with this situation being more common among schooled populations. The implicit references to cultural knowledge and customs in intelligence tests are numerous. Indeed, it could be argued that these references are impossible to avoid. It has been argued repeatedly that cross-cultural differences in intelligence test scores that are often reported are strongly influenced by the differential knowledge of the culture of the test. Individuals who share a cultural background with the test developer are more likely to obtain higher scores than are people from other cultural groups.

During recent decades, various guidelines have been formulated to describe best practices in the design of cross-cultural tests and preparing translations. For example, the International Test Commission defined a set of 22 guidelines that are useful when designing cross-cultural instruments.

Various procedures can be employed to examine the adequacy of an instrument, for example, an analysis of the contents of a test by a person (or group) with a thorough knowledge of the target culture. This person should be able to answer the question of whether the cognitive constructs of the test (e.g., memory, visuospatial skills) could be measured in this group with this instrument. In addition, various statistical techniques can be applied. These techniques require that data be collected in various cultural groups. The most frequently employed techniques address the question of whether it is reasonable to assume that the data obtained in various cultural groups can be treated statistically as coming from a single population: Can the two cultural groups be seen psychologically as belonging to the same (imaginary) statistical population? A frequently employed technique is factor analysis that examines the intertest correlations. If subtests of different tests measure entirely different psychological processes in different groups, it is very likely that test scores are interrelated in different ways across these groups. Suppose that two tests have been administered in each of two countries. Both of the tests measure reasoning in one country, whereas one test measures reasoning and the other test measures memory in the other country. It can be expected that people in the first country who score high on the first test will also score high on the second test (leading to a positive correlation between the tests), whereas a different pattern is expected in the second country because the people with high reasoning skills are not necessarily also the people with high memory skills. The intertest correlations will be higher in the first country than in the second country. So, a study of correlations of (sub)tests can yield valuable information about the psychological composition of an intelligence test in different countries.

4. FORMAL VERSUS INFORMAL STUDIES OF INTELLIGENCE

There are two research traditions in cross-cultural psychology to study the relationship between culture and intelligence: formal and informal (Table I). The formal and informal research traditions are comparable to (but not completely reducible to) the distinction between the universalist and relativist viewpoints, respectively.

In formal studies of intelligence, the scientific approach is seen as the normative model of good problem solving. There is an emphasis on the solution of formalized problems that are unlikely to be met in everyday life and on the correctness of solutions (e.g., "What is the capital of Poland?"). Nearly all intelligence tests are based on the formal tradition. Studies in this tradition have enlarged our insight on the relationship between culture and intelligence. First, administration of intelligence tests in various cultures has provided impressive evidence that the structure of intelligence is universal. The structure of cognitive abilities constitutes the universal "building blocks" that are combined by cultures in specific skill patterns useful in these cultures. (In a currently popular model of intelligence developed by Carroll in 1993, these cognitive abilities consist of fluid intelligence, crystallized intelligence, memory and learning, visual perception, auditory perception, fluency, and speed of cognitive and simple decision processes.) This finding provides a strong empirical argument against the view that all cognitive processes are culture bound. The importance of the panhuman structure of intelligence should not be underrated. It means that basic cognitive processes involved in memory are very likely to be the same across various cultures, even though

TABLE I
Differences of the Formal and Informal Traditions of the Study of Intelligence

Formal tradition	*Informal tradition*
Closed problem spaces	Open problem spaces
Deterministic problems	Probabilistic problems
Formalized artificial problems	Problems with a high ecological validity
Focus on correctness of solution	Focus on practical value of the solution (uncertainty reduction)
Context independence	Context dependence
The scientist as the model of the problem solver	The bricoleur as the model of the problem solver
Product oriented (psychometric approach) or process oriented (Piagetian approach)	Process oriented
Cross-cultural comparisons of test performances	Intracultural studies
Algorithmic solutions	Heuristic solutions
Brief assessment of many problems	In-depth assessment of a single problem
Solution requires conceptual theoretical knowledge	Solution requires procedural practical knowledge

Source. Van de Vijver and Willemsen (1993).

cultures have an impact on what kind of information is stored and retrieved. Views have been expressed that there are various kinds of intelligence across the globe and that, for example, Libyan intelligence is different from Japanese intelligence. Findings on the panhuman identity of intelligence show that the notion of multiple kinds of intelligence culturally cannot imply major differences in structure. The statement could also be interpreted to mean that Libyans and Japanese do not show the same scores on various subtests. The latter interpretation brings us to a second line of research in the formal tradition: determining and interpreting cross-cultural differences in mean scores on intelligence tests.

It is well established that cultures differ in mean scores on intelligence tests. The nature of these differences is debated, ranging from genetic differences to schooling differences. Furthermore, the quality of these tests to assess intelligence in groups for which such tests have not been developed and validated has been questioned. Although this controversy will not be solved in the near future, it is clear that the competition among these three competing explanations is unlikely to have a single winner. For instance, consider the evidence in favor of the genetic position. There are twin studies in behavior genetics in which a large proportion of individual differences in intelligence scores is found to be caused by genetic differences. Genes are nearly always found to be more important than the environment as determinants of score differences. However, the generalization of this finding to a cross-cultural context is not simple. The assessment problems discussed previously will also affect the size of the cross-cultural score differences observed. So, there is a problem of estimating the relative contributions of all relevant sources of variation. The current state of knowledge, and of the level of assessment procedures, is insufficient to deal with the question of the relative contributions in a decisive way.

A recent project by Georgas and colleagues compared data obtained with the third version of the Wechsler Intelligence Scale for Children (WISC-III) from various countries: Canada, Dutch-speaking countries (The Netherlands and Flanders), German-speaking countries (Germany, parts of Switzerland, and Austria), Greece, Japan, Lithuania, South Korea, Slovenia, Sweden, and the United States. The WISC-III test, one of the most widely employed intelligence tests for children, contains various subtests and yields an overall IQ score. The most salient finding of the study was the absence of major mean score differences on any subtest. No subtest showed a difference of more than 15 IQ points between the smallest and largest scores; in fact, the means of most countries on most subtests differed by less than 5 IQ points. The authors did not find the major differences that had been reported previously. One of the reasons for finding only small differences may be related to the sampling procedure. The samples in all countries were carefully drawn so as to represent the national populations of that age as accurately as possible, whereas older studies usually did not work with random samples. Although the mean differences of the WISC-III scales were not large, their patterning was meaningful. Countries that are richer and that spend more money on education tended to show higher scores.

The informal tradition is known by various names such as "everyday cognition," "indigenous cognition," and "practical intelligence." The bricoleur (i.e., "jack-of-all-trades") is the implicit model of problem solving. A bricoleur is not interested in a solution that is correct as defined by an external standard (which is typically the case in the formal tradition) but rather is interested in a solution that solves a practical problem (e.g., a leaking roof) using available means. Various studies have examined how individuals solve everyday cognitive problems such as mental arithmetic by shoppers who compare prices of products when both prices and weight of brands differ, planning behavior by cooks, and estimation by carpenters of how much wood is needed for making a table.

Results of research in the informal tradition show some consistent findings. Contrary to what might be expected, there is little or no relation between performances in informal tasks and those in formal tests. Street vendors who are very skillful in addition and multiplication when dealing with numbers familiar to them do not always show high levels of performance on tests of mental arithmetic that cover areas and numbers less familiar to them. Second, there is little transfer of cognitive skills from a familiar problem-solving situation to an unfamiliar one. This problem is well known in schooling; when confronted with a new problem, students do not recognize it as an application of a solution procedure they know. Third, there are remarkable differences between experts and novices regarding accuracy and strategies they use in problem-solving tasks.

5. EVERYDAY DEFINITIONS OF INTELLIGENCE

The most frequently quoted definition of intelligence is that given by Boring in 1923: "Intelligence is what an

intelligence test measures." This definition is nearly circular. A listing of the skills assessed by the subtests of an intelligence test, such as memory and reasoning, is not an informative definition of intelligence. The long tradition of theorizing and research about intelligence has not produced a satisfactory definition of the concept. Yet everyone seems to have an implicit idea of what the concept entails. During recent decades, an interesting series of studies of what laypeople mean by "intelligence" have been conducted, and many of these were done in infrequently studied groups (e.g., some rural groups in Kenya and Tanzania). In these studies, informants were asked to indicate the characteristics of both smart and dull children, to describe characteristics and behaviors of intelligent persons in their environment, and so forth. Basic constituents of Carroll's model, such as reasoning, memory, and speed, seem to be viewed as core elements of intelligence in many cultures. However, social aspects of intelligence, usually not considered to be salient aspects of the concept in psychological theories of intelligence, are important in everyday definitions. Obedience, knowing one's place in the family, knowing when to speak and when not to speak, and performing a task in the household without being asked were mentioned as examples of activities deemed to be relevant for intelligent children.

Recent developments in psychology can be seen as broadening the concept of intelligence and as being in line with this more context-informed conceptualization of intelligence. Sternberg introduced the concept of "successful intelligence," and Salovey introduced the concept of "emotional intelligence." Both of these concepts include elements that are hardly covered by classical theories, for example, creativity and empathy. The litmus test of these new tests is their power to predict important real-life criteria such as predicting which children will do well in school and which job applicant will become a good manager. The future will tell us whether predictions of real-life behavior can be improved by broadening the measures.

6. CONCLUSION

The relationship between intelligence and culture was not exactly love at first sight. Much intelligence research has been carried out from the implicit assumption that findings done in a single culture, typically a Western country, have panhuman validity. More recent studies have challenged this view. There is growing support for the view that intelligence is always situated in a specific cultural context and that cross-cultural variation is an important feature of intelligence. Yet the cross-cultural variation is not endless. There is important evidence for the universality of the basic processes. Cultures define how these building blocks are combined to produce culture-relevant skills. When differences among cultures are large, the ability patterns may show quite a bit of variation. Many examples can be found in the literature of cognitive specializations that are part of regular socialization patterns (e.g., the learning of tracking skills among the Bushmen). An important task for the future of intelligence theory and research is to combine the patterns of cross-cultural differences and similarities into one or more overarching models specifying which aspects of intelligence are influenced by culture (thereby also implicitly indicating which areas are not influenced) and how these aspects are influenced. It is only by combining models and findings of the various research traditions that the universalities and cultural specificities of intelligence can be appreciated.

See Also the Following Articles

Intelligence Assessment ■ Intelligence, Emotional ■ Intelligence in Humans

Further Reading

Berry, J. W., Poortinga, Y. H., Segall, M. H., & Dasen, P. R. (2002). *Cross-cultural psychology: Research and applications* (2nd ed.). Cambridge, UK: Cambridge University Press.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies* Cambridge, UK: Cambridge University Press.

Cole, M. (1996). *Cultural psychology: A once and future discipline*. Cambridge, MA: Harvard University Press.

Georgas, J., Weiss, L., Van de Vijver, F. J. R., & Saklofske, D. H. (Eds.). (2003). *Cultures and children's intelligence: A cross-cultural analysis of the WISC-III*. New York: Academic Press.

Greenfield, P. M. (1997). You can't take it with you: Why ability assessments don't cross cultures. *American Psychologist*, *52*, 1115–1124.

Harkness, J. A. (2003). Questionnaire translation. In J. A. Harkness, F. J. R. Van de Vijver, & P. P. Mohler (Eds.), *Cross-cultural survey methods* (pp. 35–56). New York: John Wiley.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.

Light, P., & Butterworth, G. (Eds.). (1992). *Context and cognition: Ways of learning and knowing*. London: Harvester Wheatsheaf.

Mishra, R. C. (2001). Cognition across cultures. In D. Matsumoto (Ed.), *Handbook of culture and psychology* (pp. 119–135). Oxford, UK: Oxford University Press.

Schliemann, A., Carraher, D., & Ceci, S. J. (1997). Everyday cognition. In J. W. Berry, P. R. Dasen, & T. S. Saraswathi (Eds.), *Handbook of cross-cultural psychology* (2nd ed., Vol. 2, pp. 177–216). Boston: Allyn & Bacon.

Serpell, R. (1993). *The significance of schooling: Life-journeys in an African society*. Cambridge, UK: Cambridge University Press.

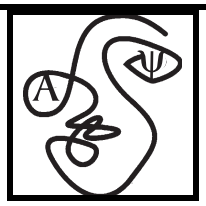
Serpell, R. (2000). Intelligence and culture. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 549–580). Cambridge, UK: Cambridge University Press.

Sternberg, R. J., & Kaufman, J. C. (1998). Human abilities. *Annual Review of Psychology*, *49*, 479–502.

Van de Vijver, F. J. R., & Leung, K. (1997). *Methods and data analysis for cross-cultural research*. Thousand Oaks, CA: Sage.

Van de Vijver, F. J. R. (1997). Meta-analysis of cross-cultural comparisons of cognitive test performance. *Journal of Cross-Cultural Psychology*, *28*, 678–709.

Van de Vijver, F. J. R., & Willemsen, M. E. (1993). Abstract thinking. In J. Altarriba (Ed.), *Culture and cognition* (pp. 317–342). Amsterdam: North-Holland.



Intelligence Assessment

Roberto Colom
Universidad Autónoma de Madrid, Madrid, Spain

1. What is Intelligence?
2. The Measurement of Intelligence
3. Intelligence Assessment Tools
4. Interpretation: Some Words
5. Recent Advances
Further Reading

GLOSSARY

cognitive ability An ability is a cognitive ability if (a) an individual's performance does not depend on any particular sensory or motor system and (b) individual differences in the ability are not correlated with sensorimotor measures.

cognitive system design An approach that includes a conceptual and procedural framework to interface item design principles with test validity; the conceptual framework broadens construct validation such as cognitive research results as supporting data, whereas the procedural framework includes several steps in which cognitive theory is integrated into test development.

g Psychometric construct resulting from the positive correlation among several diverse ability tests; *g* is what is shared or common to all of the different measures of intelligence that form a positive manifold.

intelligence General mental ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly, and learn from experience.

intelligence quotient (IQ) A score on a test that has certain characteristics, the test's raw score is usually standardized in a large representative sample, and the standardization is done within narrow age intervals, so that the standardized scores (IQ) will have the same mean and standard deviation at every age level.

In this article, intelligence is first defined and the structure of human cognitive abilities is described. Second, well-known data regarding the reliability and validity of intelligence measures are presented. Third, several remarkable intelligence assessment tools are briefly described, noting a movement from measures that are not derived from theory to measures that are theory driven. Finally, the cognitive system design approach is described as one noteworthy example of a recent advance in intelligence assessment.

1. WHAT IS INTELLIGENCE?

Intelligence is a very general mental capability involving the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly, and learn from experience. Contrary to popular belief, psychologists are nearly unanimous in their view that reasoning and problem solving are crucial landmarks of intelligent behavior.

Human cognitive abilities can be organized hierarchically on three levels or strata. On Stratum I, there are specific cognitive abilities such as general sequential reasoning, language development, associative memory, spatial relations, speech sound discrimination, ideational fluency, numerical facility, and choice reaction

©2004 Elsevier Inc.
All rights reserved.

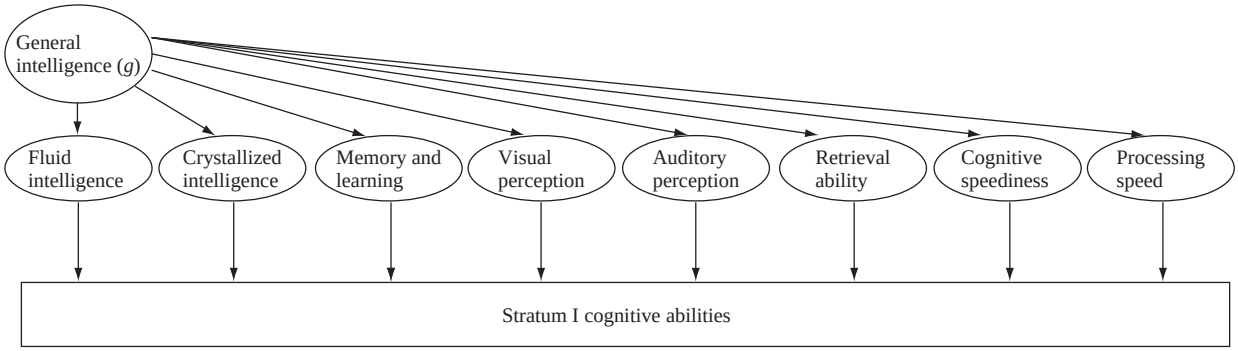


FIGURE 1 Hierarchical organization of cognitive abilities.

time. On Stratum II, there are several broad cognitive abilities such as fluid intelligence, crystallized intelligence, general memory and learning, broad visual perception, and broad auditory perception. At the top of the hierarchy, there is *g* or general intelligence. *g* accounts for half of the variability in intelligence in the general population. Figure 1 displays this hierarchical organization of cognitive abilities.

2. THE MEASUREMENT OF INTELLIGENCE

Intelligence tests are among the most reliable and valid of all psychological tests and assessments. Table I shows reliability coefficients of some of the most employed intelligence tests around the world. In addition to those reliability indexes, intelligence measured by standardized tests is the most stable psychological factor. Table II shows test–retest reliability coefficients taken from several classical longitudinal studies. These coefficients demonstrates that intelligence is a very stable human trait. Furthermore, intelligence measured by standardized tests predicts more than 60 social phenomena. Table III shows some examples.

There are different types of intelligence tests, but they all measure the same intelligence (i.e., *g*). Some

TABLE I
Reliability Coefficients of Several Intelligence Tests

Intelligence test	*Reliability coefficient*
Stanford–Binet IV	.98
WPPSI-R	.96
WISC-III	.94
WAIS-III	.98
Woodcock–Johnson III	.97
Cognitive Assessment System	.96
Kaufman Adolescent and Adult Intelligence Test	.97

TABLE II
Correlations Computed in Several Longitudinal Studies between a Test and a Retest Measure

Test (Years)	*Retest (Years)*	*Correlation coefficient*
2	15	.78
11	77	.73
14	42	.68
17	48	.83
19	61	.78
25	65	.78
30	43	.71
50	70	.90

TABLE III
Correlations between Intelligence Test Scores and Several Social Correlates

Correlate	*Correlation coefficient*
Crime	–.20
Income	.31
Health	.40
Academic success	.50
Job success	.51
Differentiation of dementia from controls	.52
Obtained level of education	.55
Job success	.63

TABLE IV
Criteria Associated with IQ Levels

IQ level	*Criteria*
40	Adult can mow lawns, do simple laundry
50	Adult can do simple carpentry, domestic work
60	Adult can repair furniture, harvest vegetables, assist electrician
75	Adult can keep small store, perform in orchestra
90	Adult can perform jobs requiring some judgment
100	Average for total population
105	Fifty–fifty chance of passing in academic high school curriculum
110	Fifty–fifty chance of graduating from college
115	Mean of freshmen in typical 4-year college
120	Mean of college graduates
130	Mean of persons receiving PhD degrees

use words or numbers and require specific cultural knowledge. Others do not and instead use shapes or designs and require knowledge of only simple universal concepts. The measurement of intelligence through a standardized intelligence quotient (IQ) test arranges people. The spread of people along the IQ continuum, from low to high, can be represented by the normal curve. Table IV shows some examples of IQ levels for various external criteria.

Psychological science has several assessment devices with high-quality standards. The remainder of this article presents some of the most remarkable intelligence tests.

3. INTELLIGENCE ASSESSMENT TOOLS

Intelligence is currently assessed through tests that are not derived from theory and through tests that are theory driven. The Wechsler scales (Wechsler Preschool and Primary Scales of Intelligence–Revised [WPPSI-R], Wechsler Intelligence Scale for Children–Third Edition [WISC-III], and Wechsler Adult Intelligence Scale–Third Edition [WAIS-III]) are examples of the former type of tests, whereas the Das–Naglieri Cognitive Assessment System (CAS) and the Woodcock–Johnson III (WJ III) are examples of the latter type of tests. Those tests can be described in terms of theoretical background, the measurement scales included, and psychometric properties such as reliability and validity.

3.1. The Wechsler Scales

David Wechsler supports the definition of intelligence as a global entity. The Wechsler scales measure intelligence from 3 to 89 years of age. The latest versions are the WPPSI-R for preschool and primary-grade children, the WISC-III for elementary and high school children, and the WAIS-III for older adolescents through old age.

3.1.1. Wechsler Preschool and Primary Scale of Intelligence–Revised

The WPPSI-R battery provides a full-scale IQ but also has separate verbal and performance scales. The verbal subtests include information, comprehension, arithmetic, vocabulary, similarities, and sentences. The performance subtests include object assembly, block design, mazes, picture completion, geometric design, and animals/pegs. The battery allows the examiner to assist the child on early items to ensure that the child understands the test requirements.

The overall verbal IQ, performance IQ, and full-scale IQ are standard scores with the mean set at 100 and the standard deviation (SD) set at 15. The range of possible WPPSI-R full-scale IQs is 41 to 160. Standard scores are provided for the separate subtests, each with a mean of 10 and an SD of 3. The battery was standardized in the United States on 1700 children from 3 years to 7 years 3 months of age and matched 1986 census bureau estimates on the stratification variables of sex, race, geographic region, parental occupation, and parental education.

The internal consistency coefficients were .95 for the verbal IQ, .92 for the performance IQ, and .96 for the full-scale IQ. The coefficients for the individual performance subtests ranged from .63 for object assembly to .85 for block design, with a median coefficient of .79. For the individual verbal subtests, the values ranged from .80 for arithmetic to .86 for similarities, with a median coefficient of .84. The test–retest coefficient for the full-scale IQ was .91 for 175 children tested twice with a time interval of 4 weeks.

The battery's manual provides several validity studies, factor-analytic results, research overviews, and interpretive tables that provide a great amount of information.

3.1.2. Wechsler Intelligence Scale for Children–Third Edition

The WISC-III battery offers standard scores on four factors: verbal comprehension, perceptual organization, freedom from distractibility, and processing speed. The first two factors consist of verbal subtests. and the latter two factors consist of performance subtests. The verbal mandatory subtests are information, similarities, arithmetic, vocabulary, and comprehension. The performance mandatory subtests are picture completion, picture arrangement, block design, object assembly, and coding. There are verbal and performance complementary subtests: digit span for verbal and mazes and symbol search for performance. Processing speed results from symbol search and coding. Freedom from distractibility results from arithmetic and digit span. The range of possible WISC-III full-scale IQs is 40 to 160.

The battery was standardized in the United States on 2200 children 6 to 16 years of age and was stratified by age, sex, race, geographic region, and parents' education. The average reliabilities across the age groups were .95 for the verbal IQ, .91 for the performance IQ, .96 for the full-scale IQ, .94 for the verbal comprehension index, .90 for the perceptual organization index, .87 for the freedom from distractibility index, and .85 for the processing speed index.

3.1.3. Wechsler Adult Intelligence Scale–Third Edition

The WAIS-III battery was designed to parallel the WISC-III. The WAIS-III consists of verbal, performance, and full-scale IQs as well as indexes on four factors. There are seven verbal subtests and seven performance subtests. Three factors have the same names as do WISC-III factors: verbal comprehension, perceptual organization, and processing speed. The fourth factor is called working memory. In addition to arithmetic and digit span, the working memory factor includes the letter–number sequencing subtest.

The WAIS-III has six regular verbal subtests and five mandatory performance subtests. The verbal subtests are vocabulary, similarities, arithmetic, digit span, information, and comprehension. The performance subtests are picture completion, picture arrangement, block design, matrix reasoning, and digit–symbol coding.

The battery was standardized in the United States on 2450 adults (selected according to 1995 census data) and was stratified according to age, sex, race, geographic region, and educational level. The participants were divided into 13 age groups, ranging from 16–17 to 85–89 years, with each group consisting of 100 to 200 people. The average split-half reliability coefficients across the age groups were .97 for the verbal IQ, .94 for the performance IQ, and .98 for the full-scale IQ. The average individual subtest reliabilities ranged from .93 for vocabulary to .70 for object assembly, with a median coefficient of .85. The stability coefficients were .95 for verbal IQ, .90 for performance IQ, and .97 for full-scale IQ (those coefficients were computed from 394 participants tested twice with a time interval of 5 weeks).

The range of possible WAIS-III full-scale IQs is 45 to 155. The WAIS-III computes scaled scores for each individual based exclusively on chronological age. The use of four indexes derives from the results of factor analysis. Several factor analyses have shown that the Wechsler performance scale measures visual–spatial ability (*Gv*), but some researchers argue that the performance scale measures a mixture of *Gv* and nonverbal reasoning (*Gf*).

3.2. Cognitive Assessment System

The CAS was developed according to the PASS (Planning, Attention, Simultaneous, Successive) theory of intelligence. The theory proposes that human cognitive functioning is based on four essential activities:

- Planning (which provides cognitive control, use of processes and knowledge, intentionality, and self-regulation to achieve a given goal)
- Attention (which provides focused, selective cognitive activity over time)
- Simultaneous processing (one form of operating on information)
- Successive processing (another form of operating on information)

The subtests are organized into four scales. Planning subtests require devising, selecting, and using efficient plans of action to solve the problems, regulate the effectiveness of the plans, and self-correct if necessary. Attention subtests require attending selectively to a particular stimulus and inhibiting attending to distracting stimuli. Simultaneous processing subtests require integrating stimuli into groups. Successive processing subtests require integrating stimuli in their serial order or appreciating the linearity of stimuli.

The CAS yields scores for the planning, attention, simultaneous, successive, and full scales, with a

normative mean of 100 and an SD of 15. All subtests are set at a normative mean of 10 and an SD of 3.

The battery was standardized in the United States on 2200 participants 5 to 17 years of age and was stratified by age, sex, race, geographic region, educational placement, and parents' education according to census reports and matches population characteristics in the variables used.

The internal consistency reliability estimates for the CAS were as follows: full scale = .96, planning and attention = .88, and simultaneous processing and successive processing = .93. The test–retest reliabilities for the CAS were as follows: full scale = .91, planning = .85, attention = .82, simultaneous processing = .81, and successive processing = .86. A total of 215 participants were tested twice, with a time interval of 21 days.

3.3. Kaufman Adolescent and Adult Intelligence Test

The Kaufman Adolescent and Adult Intelligence Test (KAIT) is based mainly on the Horn–Cattell theory of intelligence. Fluid intelligence (*Gf*) involves culture-fair novel tasks and taps problem-solving skills and the ability to learn. Crystallized intelligence (*Gc*) refers to acquired skill, knowledge, and judgments learned within a given society.

Thus, the battery provides fluid, crystallized, and composite IQs, with a mean of 100 and an SD of 15. It consists of a core battery of six subtests and an expanded battery. Each subtest yields age-based scores with a mean of 10 and an SD of 3. Crystallized intelligence measures the acquisition of facts and problem-solving ability using stimuli derived from schooling and cultural experiences. Fluid intelligence measures a person's adaptability when solving new problems. The KAIT fluid subtests emphasize reasoning rather than visual–spatial ability (contrary to the Wechsler performance scale).

The KAIT was standardized in the United States on 2000 participants 11 to 94 years of age and was stratified according to sex, race, geographic location, and socioeconomic status. The reliability coefficients were .95 for *Gc* and *Gf* and .97 for composite IQ. Test–retest reliabilities were .94 for *Gc* and composite IQ and .87 for *Gf*. A total of 153 participants were tested twice, with a time interval of 1 month.

Both exploratory and confirmatory factor analyses support the construct validity of the KAIT. Moreover, a correlation of .82 was observed between WISC-R full-scale IQ and the KAIT composite IQ, and a correlation of .85 was observed between WAIS-R full-scale IQ and the KAIT composite IQ.

3.4. Woodcock–Johnson III

The WJ III consists of two distinct batteries: the WJ III Tests of Cognitive Abilities and the WJ III Tests of Achievement. Together, these batteries comprise a wide age range system for measuring general intellectual ability (*g*), specific cognitive abilities, oral language, and academic achievement. This subsection reviews only the WJ III Tests of Cognitive Abilities, where the standard battery and the extended battery can be found.

The standard battery includes verbal comprehension, visual–auditory learning, spatial relations, sound blending, concept formation, visual matching, numbers reversed, incomplete words, auditory working memory, and delayed visual–auditory learning. The extended battery includes general information, retrieval fluency, picture recognition, auditory attention, analysis–synthesis, decision speed, memory for words, rapid picture naming, planning, and pair cancellation.

The WJ III is based on current theory and research on the structure of human cognitive abilities. The theoretical basis is derived from the Cattell–Horn–Carroll theory of cognitive abilities (Fig. 1). The WJ III is a measurement model of this theory, and the design criteria place emphasis on providing the greatest practical breadth in Stratum II abilities: comprehension knowledge, long-term retrieval, visual–spatial thinking, auditory processing, fluid reasoning, processing speed, and short-term memory. The measured abilities are as follows:

- *Gf* (fluid intelligence) through concept formation (induction), analysis–synthesis (general sequential deductive reasoning), and planning (general sequential reasoning)
- *Gc* (crystallized intelligence) through verbal comprehension (lexical knowledge and language development) and general information (general verbal information)
- *Glr* (long-term retrieval) through visual–auditory learning (associative memory), delayed visual–auditory learning (associative memory), and retrieval fluency (ideational fluency)
- *Gv* (visual–spatial thinking) through spatial relations (visualization and spatial relations), picture recognition (visual memory), and planning (spatial scanning)
- *Ga* (auditory processing) through sound blending (phonetic coding synthesis), incomplete words

(phonetic coding analysis), and auditory attention (speech sound discrimination and resistance to auditory stimulus distortion)
- *Gs* (processing speed) through visual matching (perceptual speed), decision speed (semantic processing speed), rapid picture naming (naming facility), and pair cancellation (attention and concentration)
- *Gsm* (short-term memory) through numbers reversed (working memory), auditory working memory (working memory), and memory for words (memory span)

The data for WJ III norms were collected from a large, nationally representative sample of 8818 participants gathered in more than 100 geographically diverse U.S. communities. The preschool sample (2–5 years of age) was composed of 1143 participants. The kindergarten through 12th-grade sample was composed of 4783 participants. The college/university sample consisted of 1165 undergraduate and graduate students. The adult sample was composed of 1843 participants. The norming sample was selected to be representative of the U.S. population from 24 months to 90 years of age or over. The participants were randomly selected within a stratified sampling design that controlled for census region, community size, sex, race, type of school, type of college/university, education of adults, occupational status of adults, and occupation of adults in the labor force.

The WJ III general intellectual ability score is a general intelligence (*g*) score. It represents the first principal component obtained from principal component analyses. Using weights based on principal component analyses means that all subtest weights are optimal. In contrast, tests such as the Wechsler intelligence scales weight all subtests equally. Using principal component analyses of cognitive measures as a basis for prescribing different test weights gives the best statistical estimate of general intelligence.

Reliability coefficients ranged from .74 for planning to .97 for rapid picture naming. The construct validity of WJ III was tested by confirmatory factor analyses. The breadth of abilities measured by the WJ III is described by *g* plus several broad cognitive abilities plus several narrow cognitive abilities. Nearly all tests load exclusively on a single factor, suggesting that the cognitive tests have minimized the influence of construct-irrelevant variance.

The WJ III general intellectual ability scores had correlations ranging from .67 to .76, with composite scores from the WPPSI-R, the WISC-R, or the Stanford–Binet IV. The WJ III matches its intended theory closely. It is a good translation of theory into practice.

4. INTERPRETATION: SOME WORDS

The examiner must produce hypotheses about an individual's strengths and weaknesses based on theory, research knowledge, and clinical ability. The survival hypotheses must form practical recommendations.

Intelligence assessment devices must be selected prudently. These devices must meet sound psychometric standards, but the examiner should keep in mind that each examinee has unique characteristics. Alan Kaufman articulated five basic principles for intelligence assessment:

1. The items on an intelligence test measure what the individual has learned.
2. The subtests included on an intelligence test are nonexhaustive samples of behavior.
3. The items on an intelligence test assess mental functioning under fixed conditions.
4. The subtests on an intelligence test are useful when interpreted within an information processing model: how information enters the system, how it is interpreted and processed, how it is stored for later retrieval, and how it is expressed.
5. The hypotheses produced from an intelligence assessment should be supported from several sources: the person's behavior during test administration, the pattern of responses across several subtests, background information from parents, teachers, and/or peers, and so forth.

These basic principles are appealing, but the examiner must be cautious because there is always the temptation to neglect objective data in the face of other sources of much more subjective evidence. Therefore, the examiner must consider asking some colleagues when the distance between the objective result and the subjective appreciation increases.

5. RECENT ADVANCES

Intelligence assessment is moving from an exclusively empirically rooted test construction to a more

theoretically based one. The WJ III and the KAIT are some examples. But there are other instances as well. The following focuses briefly on the recent applied research development known as the cognitive system design (CSD) approach.

The CSD approach emphasizes manipulating construct validity by the cognitive complexity of the items. These items can be designed to measure targeted aspects of processing by detailed specifications of the stimulus features that influence processing. The stages are as follows:

- Specify general goals of measurement.
- Identify design features in task domain (task-general features [e.g., model, format, conditions] and task-specific features).
- Develop a cognitive model (review theories, select or develop model for psychometric domain, revise model, and test model).
- Evaluate cognitive model for psychometric potential (evaluate cognitive model plausibility on current test, evaluate impact of complexity factors on psychometric properties, and anticipate properties of new test).
- Specify item distributions on cognitive complexity (distribution of item complexity parameters and distribution of item features).
- Generate items to fit specifications.
- Evaluate cognitive and psychometric properties for revised tests domain (estimate component latent trait model parameters, estimate plausibility of cognitive model, evaluate impact of complexity factors on psychometric properties, evaluate plausibility of the psychometric model, and calibrate final item parameters and ability distributions).
- Conduct a psychometric evaluation (measuring processing abilities and bank items by cognitive processing demands).
- Assemble test forms to represent specifications (fixed content test and adaptive test).
- Validate.

Cognitive item response theory (IRT) models are mathematical models of cognitive processes and IRT models of response patterns. These models contain parameters to represent the cognitive demands of items as well as the person's ability.

Susan Embretson has applied the CSD approach to develop a test of abstract reasoning called ART. ART contains matrix problems very similar to those of the Advanced Progressive Matrices (APM) test. A bank of 150 items was generated analytically from a list of objects and attributes. Empirical results supported the CSD approach to test development. Working memory capacity was the primary cognitive aspect of item solving. Memory load predicted item difficulty. Items were banked by their cognitive demands by use of the memory load scores from the cognitive model. Moreover, ART scores were indistinguishable from APM scores. ART achieved the same internal consistency and factorial validity as did APM but from fewer items, suggesting that the CSD approach may eliminate features from items that are irrelevant to abstract reasoning.

Designing intelligence tests from cognitive theory is now feasible. Good results may be expected for a broad array of items if a high-quality cognitive theory about the item stimulus features that influence processing is developed. Future testing may change if the generating potential of cognitive theory is realized. The potential size of the generable items bank is quite large. Even item construction could be fully computerized for future tests. The test of the future will be a set of generating principles with previously known relationships to people's performance. This obviously requires a fully developed cognitive model.

See Also the Following Articles

Clinical Assessment ■ Intelligence and Culture ■ Intelligence, Emotional ■ Intelligence in Humans ■ Psychological Assessment, Standards and Guidelines for ■ Psychometric Tests ■ Psychophysiological Assessment

Further Reading

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies*. Cambridge, UK: Cambridge University Press.

Colom, R. (2003). Multiple cognitive abilities assessment. In R. Fernández-Ballesteros (Ed.), *Encyclopedia of psychological assessment*. Thousand Oaks, CA: Sage.

Colom, R., & Andrés-Pueyo, A. (2000). The study of human intelligence: A review at the turn of the century. *Psychology in Spain*, *4*(1), 167–182.

Fernández-Ballesteros, R., & Colom, R. (2004). The psychology of human intelligence in Spain. In R. J. Sternberg (Ed.), *International handbook of the psychology of human intelligence* (pp. 79–103). Cambridge, UK: Cambridge University Press.

Horn, J. (1989). Cognitive diversity: A framework for learning. In P. Ackerman, R. Sternberg, & R. Glaser (Eds.), *Learning and individual differences*. New York: Freeman.

Kaufman, A. S., & Kaufman, N. L. (1993). *Manual for Kaufman Adolescent and Adult Intelligence Test (KAIT)*. Circle Pines, MN: American Guidance Service.

McGrew, K. S., & Woodcock, R. W. (2001). *Technical manual: Woodcock–Johnson III*. Chicago: Riverside.

Naglieri, J. A., & Das, J. P. (1997). *Das–Naglieri Cognitive Assessment System*. Chicago: Riverside.

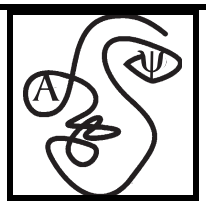
Neisser, U., Boodoo, G., Bouchard, T., Boykin, A., Brody, N., Ceci, S., Halpern, D., Loehlin, J., Perloff, R., Sternberg, R., & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, *51*(2), 77–101.

Van der Linden, W., & Hambleton, R. (1997). *Handbook of modern item response theory*. New York: Springer.

Wechsler, D. (1989). *Manual for the Wechsler Preschool and Primary Scale of Intelligence–Revised (WPPSI-R)*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1991). *Manual for the Wechsler Intelligence Scale for Children–Third Edition (WISC-III)*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1997). *Manual for the Wechsler Adult Intelligence Scale–Third Edition (WAIS-III)*. San Antonio, TX: Psychological Corporation.



Intelligence, Emotional

Cary Cherniss

Rutgers University, Piscataway, New Jersey, USA

1. Why the Interest in Emotional Intelligence?
2. Historical Roots of and Precursors to the Concept
3. Various Models of Emotional Intelligence
4. Measurement of Emotional Intelligence
5. The Relationship between Emotional Intelligence and Personality
6. The Contribution of Emotional Intelligence to Health, Work, and Educational Adjustment
7. Can Emotional Intelligence Be Taught?
Further Reading

GLOSSARY

cognitive abilities Higher order mental faculties that involve the acquisition, analysis, and categorization of information as well as reasoning and problem solving.

competency A learned capability that contributes to superior performance.

conative abilities Having to do with motivation as opposed to emotion or cognition.

360-degree assessment An assessment in which a number of people who know an individual answer questions about that individual's attitudes, ability, or behavior; in the workplace, the raters are usually the individual's supervisors, peers, and subordinates.

tipping point A point in the linear relationship between two variables at which an additional small increase in the first variable results in a sharp increase in the second variable.

Emotional intelligence refers to a set of abilities that involve the way in which people perceive, express, understand, and manage their own emotions as well as the emotions of others. Although there are several different definitions and models found in the literature, there is considerable overlap among them. At the core of every definition are the abilities to perceive emotion clearly in oneself and in others, to use emotions to facilitate thought and action, to understand how emotions affect one's own behavior and that of others, and to regulate one's own emotional reactions.

1. WHY THE INTEREST IN EMOTIONAL INTELLIGENCE?

For more than 100 years, psychologists have been trying to identify, measure, and modify those aspects of personality and behavior that most strongly affect a person's ability to adapt successfully to the demands of living. Many of these efforts have focused on cognitive ability, that is, the way in which one acquires, codes, stores, and retrieves information about the environment. However, no matter how one measures cognitive ability, the amount of variability in outcomes attributable to it always proves to be rather limited. Estimates vary depending on the nature of the task and context, but cognitive ability, as measured by intelligence quotient (IQ) tests or other tests of general mental ability, usually accounts for approximately 10% of the variability in performance outcomes—and sometimes much

©2004 Elsevier Inc.
All rights reserved.

less. Up to now, the largest percentage attributable to cognitive ability has been approximately 25%. In addition, many tests of cognitive ability prove to be culturally biased; thus, the use of such tests perpetuates discrimination against many groups within society.

The significant but ultimately limited role of cognitive ability alone in helping people to adapt successfully has led to several attempts to identify other kinds of abilities that contribute to effective performance in school, interpersonal relationships, and the workplace. Emotional intelligence (EI) represents one such set of abilities.

2. HISTORICAL ROOTS OF AND PRECURSORS TO THE CONCEPT

Even before psychologists began to measure intelligence, Darwin wrote about the adaptive importance of emotion in human and animal life. In 1872, he published a book in which he proposed that emotions serve two important survival and adaptive functions. First, they energize and motivate activity. Fear, to take an obvious example, helps one to be more vigilant and, when danger is spotted, to run faster from it. Second, emotions also serve a signal function, conveying important information from one animal in a group to others. If an elk, grazing in a herd, notices a mountain lion off in the distance, the elk's fear reactions alert the rest of the herd, and the entire herd quickly sets off in flight, helping the elk to survive and pass on their genes to the next generation.

As early as 1920, some of the pioneers in the study and measurement of intelligence recognized the need to move beyond purely cognitive-based abilities in their conceptions and to include other kinds of abilities. Thorndike, for instance, wrote about "social intelligence" in an article published in 1920. In 1943, Wechsler wrote that "nonintellective" factors or "affective and conative abilities" should be considered part of general intelligence, based on their capacity to facilitate intelligent behavior. During the early 1980s, Gardner developed his concept of "multiple intelligences," which included "intrapersonal" and "interpersonal" intelligences as well as the more traditional cognitive-based types. Also during the early 1980s, Bar-On began to study what he later referred to as "emotional and social intelligence." Around the same time, Sternberg began to write about "practical intelligence." In 1990, Salovey and Mayer published an article in which they used the term "emotional intelligence" in the title, and they subsequently began developing their own model and measures of the construct.

There was another stream of activity in psychology that was a relevant part of the concept's history, even though it was not explicitly linked to intelligence. For decades, personality and social psychologists had been studying personal attributes that seemed to help people be more effective in a variety of contexts, and some of these attributes involved one's ability to perceive, express, and/or manage emotions in oneself and in others. For instance, there was research on empathy, defined as the ability to sense what others are feeling and to take their perspective, suggesting that both children and adults who scored higher in this quality tended to be more successful in both vocational and academic contexts. Similarly, there was research suggesting that the ability to express emotion in a clear and compelling way could have a positive impact on individual and group performance in a variety of contexts. Moreover, clinical work on anger and anger management suggested that awareness of one's emotional reactions and what triggers them can greatly influence one's ability to regulate emotion, which in turn affects well-being and success. Current research and applications relating to EI draw on all of these historical roots.

3. VARIOUS MODELS OF EMOTIONAL INTELLIGENCE

Several different models of EI have emerged since 1990, and such diversity has contributed to a certain amount of controversy. Although there are important differences among the models, there also are many similarities. In fact, there probably is as much consensus about the EI construct as there is about the concept of general intelligence, a topic that has certainly seen its own share of conflict and controversy.

Salovey and Mayer called their model an "ability" model. Using a deductive approach, they identified four "branches" that are related in a hierarchical way: (a) the ability to perceive emotions accurately, (b) the ability to use emotions to facilitate thought, (c) the ability to understand emotions, and (d) the ability to manage emotions.

A second popular model is Goleman's "performance-based" model. Like Salovey and Mayer's model, the first component of Goleman's model involves perception of emotion. In this case, however, the focus is on awareness

of one's own emotions (Self-Awareness). A second component of the model involves awareness of emotions in others (Social Awareness). These first two components involve recognition of emotion. The other two components involve regulation of emotion: Self-Management and Relationship Management. Also like Salovey and Mayer, Goleman proposed a kind of hierarchical relationship among these components, with Self-Awareness constituting the foundation for Self-Management and Social Awareness, which in turn provides a foundation for Relationship Management. Goleman added another level to his model, consisting of approximately 20 "competencies" that are linked to the four basic dimensions of EI. For instance, under the Self-Management dimension, there are competencies such as flexibility and initiative. Goleman stressed that these competencies are not synonymous with EI, although EI provides the foundation for them. They are learned capabilities that contribute to superior performance.

Another model that has attracted considerable attention is Bar-On's model, which also is a competency-based model. Bar-On's model is composed of five dimensions: (a) Intrapersonal, which includes competencies such as emotional self-awareness and assertiveness; (b) Interpersonal, which includes empathy and interpersonal relationships, among others; (c) Stress Management, which incorporates stress tolerance and impulse control; (d) Adaptability, which consists of flexibility, reality testing, and problem solving; and (d) General Mood, covering optimism and happiness.

One reason why the models differ is that they were developed in different ways and for different purposes. Salovey and Mayer's model was developed deductively. Those researchers began with a general conception of EI and then defined their components based on that conception. Goleman, on the other hand, developed his model inductively. He identified the competencies that previous research suggested are most predictive of superior performance and then arranged them in clusters that seemed closely related to the basic components of EI. Bar-On's approach also was more inductive. Beginning with clinical work on life adjustment, he identified a set of skills that seemed to help people cope with demands and pressures. Then, using factor-analytic procedures, he ended up with his model of five major dimensions and 15 competencies.

Despite these differences in the models, a careful comparison suggests that there is considerable overlap. For instance, Salovey and Mayer's Perception of Emotion and Understanding Emotion dimensions seem to be similar to Goleman's Self-Awareness and Social Awareness dimensions as well as to the emotional self-awareness and empathy subscales of Bar-On's model. Similarly, Salovey and Mayer's Emotional Facilitation of Thinking and Managing Emotions dimensions seem to be similar to Goleman's Self-Management and Relationship Management dimensions as well as to the Interpersonal, Stress Management, and Adaptability dimensions of Bar-On's model. Thus, although there are important differences among the three models, there also are many similarities.

4. MEASUREMENT OF EMOTIONAL INTELLIGENCE

Measures for various aspects of EI have existed for some time. For instance, there are measures of empathy that can be used to assess one's ability to perceive clearly another person's mood or emotional response. Similarly, there are measures of emotional self-awareness and self-regulation. However, during the past decade or so, there also has been a concerted attempt to develop and refine instruments explicitly designed to measure all of the major facets of EI.

The first such instrument is Bar-On's Emotional Quotient Inventory (EQ-i). The EQ-i is a self-assessment instrument that asks an individual to answer a series of questions about his or her own emotional and social competence based on the way in which the respondent feels, thinks, and behaves most of the time. Bar-On originally developed the instrument during the 1980s and has done considerable research on it over the years. Plake and Impara's critical review in 2002 suggested that the EQ-i has good reliability and validity.

Another instrument is the Mayer, Salovey, and Caruso Emotional Intelligence Test (MSCEIT), which is an ability test in which an individual performs a number of tasks designed to test various aspects of EI. For instance, in one section of the test, the individual is shown a series of faces and must rate how much sadness, anger, jealousy, and the like are displayed in each one. This section is designed to assess the Perception of Emotion dimension. The MSCEIT is the latest in a series of tests that this research group has developed; thus, it has less psychometric research behind it at this point. But a study by Mayer, Salovey and Caruso provided some promising support for the instrument's reliability and validity.

A third test that sometimes is used to measure EI is the Emotional Competence Inventory (ECI), developed by Boyatzis and Goleman. The original version of this test

was a "360-degree assessment" in which a number of people who know an individual answer questions about that person's social and emotional competence. Many of the items are based on earlier research that identified the competencies most important for success in the workplace. Like the MSCEIT, the ECI is relatively new, but the body of research reported by Sala in 2002 suggests that the instrument has good reliability and validity.

There have been several other attempts to develop measures of EI, but these have received less attention than have the EQ-i, MSCEIT, and ECI. Also, there has been less psychometric research reported on those other measures.

5. THE RELATIONSHIP BETWEEN EMOTIONAL INTELLIGENCE AND PERSONALITY

One of the controversies surrounding the concept of EI concerns whether it is simply a new name for a set of personality traits that have been studied often in the past. Empirical research has addressed this issue by examining the strength of the relationship between common measures of EI and various personality traits. Most attention has focused on the relationship between EI and the Five-Factor Model of personality. A meta-analysis by Van Rooy and Viswesvaran, based on 69 studies involving more than 12,000 individuals, found that although there is some overlap, the relationships generally tend to be modest. The highest correlate of EI among the "Big Five" factors was Extraversion (r = .34). This same study also found that EI demonstrated significant incremental validity over the Big Five factors but that the Big Five did not add to EI, suggesting that EI is a better predictor of performance than is personality.

6. THE CONTRIBUTION OF EMOTIONAL INTELLIGENCE TO HEALTH, WORK, AND EDUCATIONAL ADJUSTMENT

Although there is some controversy about how much EI contributes to health, work, and educational outcomes, there is considerable research showing that it does have an impact. The research comes from two sources. First, there are recent studies that have used one of the explicit measures of EI that have appeared during the past decade or so. Van Rooy and Viswesvaran's meta-analysis found that, across all studies and measures, the correlation between EI and work performance was .24. For academic performance, the relationship was weaker; the overall correlation was .10. As is often the case with meta-analyses, the strength of the relationship varied considerably depending on the context of the study and the measures used. For instance, studies using an early measure of EI developed by Salovey and Mayer found a correlation of .32 between EI and performance, whereas researchers using the EQ-i published canonical correlations between EI and occupational performance as high as .74. Some researchers also have reported stronger associations between EI and academic performance. For example, one study found a correlation of .41 between EQ-i scores and grade point average in a group of American high school students.

Another source of research on the link between EI and important outcomes are the countless studies that have examined specific facets of EI during the past half-century or so. Such research has shown, for instance, that awareness of one's own emotional reactions plays a vital role in stress regulation, anger management, and other processes that affect both physical and behavioral health. Emotional self-regulation plays an even greater role in health outcomes and also is a strong predictor of success in school and the workplace. Similarly, empathy (i.e., the ability to perceive emotions in others) has been shown to predict academic and vocational adjustment and to provide the foundation for positive social relationships that contribute to physical and behavioral health in various ways.

Most of the research on the link between EI and important outcomes probably provides an underestimation of EI's significance because researchers typically look only at simple linear relationships. There is some research and theory suggesting that the relationship between EI and health or performance is more complex. For instance, Goleman suggested that EI is especially important in determining which individuals, among those occupying a particular role in an organization or occupation, will become superior performers or "stars." Research on competence in work organizations summarized by McClelland in 1998 tends to support this view. Also, McClelland and his students showed that there tend to be "tipping points" in the relationship between emotional competencies and performance. This means that up to a certain level of competence, increases in competence do not result in significant increases in performance; however, once the level of competence passes a tipping point, there is a sharp increase in performance. Such complex relationships

could tend to suppress the strength of simple correlational and regression estimates.

7. CAN EMOTIONAL INTELLIGENCE BE TAUGHT?

At least since Freud's early work in psychoanalysis, psychotherapists of various kinds have been teaching their patients to become more aware of their emotional responses, to develop greater understanding of their emotions, to perceive the feelings of others more accurately, and to regulate their own emotions more effectively. Similarly, since the beginning of civilization, parents and teachers have been helping children to better understand and control their emotions and those of others. If these efforts had not been effective, civilization as we know it probably could not exist. Thus, it seems obvious that EI can be taught.

However, less obvious is the question of how EI can best be taught and how much change an intentional intervention can actually bring about in children or adults. The evidence suggests that EI can be taught but that the process is more time-consuming and difficult than many self-appointed gurus, teachers, and trainers often suggest. Perhaps the clearest indication that EI can be taught comes from efforts to help people better manage their emotions. Recent research on stress and anger management programs suggests that methods are now available to help many people who have problems with self-regulation of emotion. There also have been experimental procedures developed to help people become more aware of others' emotions and to identify those emotions more accurately. Moreover, a number of applied psychologists have developed educational curricula and teaching methods that seem to be effective in helping many children to develop social and emotional competence.

Unfortunately, all of these intervention strategies require a highly favorable context to be effective for many individuals. First, and most important, the individual usually must be highly motivated and willing to devote considerable time and effort to the change effort. Second, the teacher or trainer must have a high degree of skill, particularly in the social and emotional domain, to be helpful. Third, the social environment needs to provide support and encouragement of various kinds. When all of these critical ingredients are present, well-designed and -implemented interventions can help children and adults to become more emotionally intelligent.

See Also the Following Articles

Emotion ■ Intelligence and Culture ■ Intelligence Assessment ■ Intelligence in Humans ■ Personality Assessment ■ Traits ■ Work Adjustment

Further Reading

Bar-On, R., & Parker, J. D. A. (2000). *The handbook of emotional intelligence*. San Francisco: Jossey–Bass.

Cherniss, C., & Goleman, D. (2001). *The emotionally intelligent workplace*. San Francisco: Jossey–Bass.

Ciarrochi, J., Forgas, J. P., & Mayer, J. D. (2001). *Emotional intelligence in everyday life: A scientific inquiry*. Philadelphia: Taylor & Francis.

Goleman, D. (1995). *Emotional intelligence*. New York: Bantam.

Matthews, G., Zeidner, M., & Roberts, R. D. (2002). *Emotional intelligence: Science and myth*. Cambridge, MA: MIT Press.

Mayer, J. D., Salovey, P., & Caruso, D. R. (in press). Emotional intelligence: Theory, findings, and implications. *Psychological Inquiry*.

McClelland, D. C. (1998). Identifying competencies with behavioral-event interviews. *Psychological Science*, *9*, 331–339.

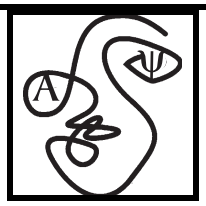
Plake, B. S., & Impara, J. C. (Eds.). (2002). *The fourteenth Buros mental measurement yearbook*. Lincoln, NE: Buros Institute for Mental Measurement.

Sala, F. (2002). *Emotional Competence Inventory technical manual*. Boston: Hay Group.

Salovey, P., & Mayer, J. (1990). Emotional intelligence. *Imagination, Cognition, and Personality*, *9*, 185–211.

Van Rooy, D. L., & Viswesvaran, C. (in press). Emotional intelligence: A meta-analytic investigation of predictive validity and nomological net. *Journal of Vocational Behavior*.

Wagner, R. K. (1997). Intelligence, training, and employment. *American Psychologist*, *52*, 1059–1069.



Intelligence in Humans

Robert J. Sternberg
Yale University, New Haven, Connecticut, USA

1. Theories of Intelligence
2. Development of Intelligence
3. Measuring Intelligence
4. The Heritability and Malleability of Intelligence
Further Reading

GLOSSARY

crystallized intelligence One of two major subfactors of general intelligence; it represents the accumulation of knowledge over the life span of the individual and may be measured by tests in areas such as vocabulary, general information, and achievement.

deviation IQs A means of determining intelligence test scores, based on deviations from an average score, calculated such that the normative equivalent for the median score is 100, approximately 68% of the scores are computed to fall between 85 and 115, and approximately 95% of the scores are computed to fall between 70 and 130.

emotional intelligence The ability to perceive and express emotion, assimilate emotion in thought, understand and reason with emotion, and regulate emotion in the self and others.

fluid intelligence One of two major subfactors of general intelligence; it represents the acquisition of new information or the grasping of new relationships and abstractions about known information.

heritability coefficient The degree to which heredity contributes to individual differences in intelligence, expressed in terms of a number on a scale from 0 to 1 such that a coefficient of 0 means that heredity has no influence on variation among people, whereas a coefficient of 1 means that heredity is the only influence on such variation.

intelligence The ability to implement goal-directed adaptive behavior.

mental age A means of indicating a person's level of intelligence (generally in reference to a child), based on the individual's performance on tests of intelligence, by indicating the chronological age of persons who typically perform at the same level of intelligence as the test taker.

mental retardation Low level of intelligence, usually reflected by both poor performance on tests of intelligence and poor adaptive competence, that is, the degree to which a person functions effectively within a normal situational context.

ratio IQ A means of indicating performance on intelligence tests, based on a quotient of mental age divided by chronological age multiplied by 100.

reaction range The broad limits within which a particular attribute (e.g., intelligence) may be expressed in various possible ways given the inherited potential for expression of the attribute in the particular individual.

theory of multiple intelligences A theory suggesting that intelligence is composed of eight distinct constructs that function somewhat independently but may interact to produce intelligent behavior: bodily–kinesthetic intelligence, interpersonal intelligence, intrapersonal intelligence, linguistic intelligence, mathematical–logical intelligence, musical intelligence, naturalist intelligence, and spatial intelligence.

triarchic theory of successful intelligence A theory of intelligence asserting that intelligence is composed of three aspects that deal with the relationship of intelligence to the internal world, to experience, and to the external world.

©2004 Elsevier Inc.
All rights reserved.

Intelligence is the ability to implement goal-directed adaptive behavior. In this article, the concept of intelligence is humans is examined. The article opens with a review of theories of intelligence, then discusses development of intelligence, next measuring intelligence, and finally, the heritability and malleability of intelligence.

1. THEORIES OF INTELLIGENCE

1.1. Implicit Theories

How do psychologists know what intelligence is? One way in which they find out is to ask people. For example, they might ask experts. In a 1921 symposium on the definition of intelligence, an American psychologist, Lewis M. Terman, emphasized the ability to think abstractly. Another American psychologist, Edward L. Thorndike, emphasized learning and the ability to give good responses to questions. In a similar 1986 symposium, however, psychologists generally agreed on the importance of adaptation to the environment as the key to understanding both what intelligence is and what it does. They also emphasized learning skills and understanding one's own cognitive processes.

In a set of studies published in 1981, Sternberg and colleagues asked laypeople in the United States what they thought intelligence was. Three factors emerged from the responses: practical problem solving, verbal ability, and social competence. However, the conceptions that arise depend on who is asked. In a comparable set of studies carried out in Taiwan, five factors emerged: cognitive skills, getting along with other people, understanding oneself, knowing when to show one is smart, and knowing when not to show one is smart. In other studies conducted in Kenya, cognitive skills also were emphasized less than in the West, whereas obedience, respect, and understanding people were emphasized more.

Most theories of intelligence are explicit rather than implicit; that is, they are elicited not by asking people what they mean by intelligence but rather by having people perform tasks alleged to require intelligence and then analyzing the way in which people perform. There are several different kinds of explicit theories.

1.2. Psychometric Theories

Psychometric theories have generally sought to understand the structure of intelligence: What form does it take, and what are its parts (if any)? Such theories have generally been based on and tested by the use of data obtained from tests of mental abilities. These tests include items such as vocabulary, completing series of numbers, seeing analogies between words, and visualizing what forms would look like if they were rotated in space.

The first of the major psychometric theories was that of the British psychologist Charles E. Spearman. In a 1904 article, Spearman argued that just two kinds of factors underlie all individual differences in test scores. Spearman called the first and more important kind of factor the "general factor" or *g*. It was said to pervade performance on all tasks requiring intelligence. The second kind of factor was specific to each test. But what exactly is *g*? In 1927, Spearman proposed that it might be something he labeled "mental energy."

An American psychologist, L. L. Thurstone, suggested instead that seven factors, or "primary mental abilities," pervade test performance: verbal comprehension (knowledge of vocabulary and in reading), verbal fluency (writing and producing words), number (solving simple arithmetical computation and reasoning problems), spatial visualization (mentally visualizing and manipulating objects), inductive reasoning (completing a number series or predicting the future based on past experience), memory (remembering people's names or faces), and perceptual speed (rapidly proofreading to discover typographical errors in a typed text).

Other psychologists, including Raymond B. Cattell and John B. Carroll, suggested that abilities are hierarchical. At the top of the hierarchy is *g* or general ability. But below *g* in the hierarchy are successive levels of gradually narrowing abilities, ending with Spearman's specific abilities.

Cattell suggested that general ability can be divided into two broad kinds of abilities: fluid and crystallized. Fluid abilities are the reasoning and problem-solving abilities measured by tests such as the analogies, classifications, and series completions. Crystallized abilities can be said to derive from fluid abilities and are viewed as their products, which would include vocabulary, general information, and knowledge about specific fields. John L. Horn suggested that crystallized ability more or less increases over the life span, whereas fluid ability increases during the earlier years and decreases during the later years.

Carroll proposed a "three-stratum" model of intelligence that is considered by some to be the most definitive psychometric model of intelligence because it is based on reanalysis of hundreds of data sets. According to this model, general ability is at the top of a hierarchy

of abilities. At the next lower stratum are various broad abilities such as learning and memory processes and the effortless production of many ideas. At the bottom of the hierarchy are many narrow specific abilities such as spelling ability and reasoning speed.

J. P. Guilford, an American psychologist, proposed a "structure of intellect" theory that, in its earlier versions, postulated 120 abilities. For example, in an influential 1967 work, Guilford argued that abilities can be divided into five kinds of operations, four kinds of contents, and six kinds of products. These various facets of intelligence combine multiplicatively for a total of $5 \times 4 \times 6$ or 120 separate abilities. An example of such an ability would be cognition (operation) of semantic (content) relations (product), which would be involved in recognizing the relation between lawyer and client in the following analogy problem: lawyer: client: doctor:? In 1984, Guilford increased the number of abilities proposed by his theory, raising the total to 150.

1.3. Cognitive Theories

In a 1957 address to the American Psychological Association, Lee J. Cronbach proposed that psychologists unite the "two disciplines of scientific psychology"—experimental and differential (i.e., the study of individual differences). His proposal led to cognitive theories of intelligence that are derived by experimental means.

Underlying most cognitive approaches to intelligence is the assumption that intelligence is composed of a set of mental representations (e.g., propositions, images) of information and a set of mental processes that can operate on the representations. A more intelligent person is assumed to mentally represent information better, and also to operate more quickly on these representations, compared with a less intelligent person.

In 1973, Earl B. Hunt, Nancy Frost, and Clifford E. Lunneborg showed that a critical ability underlying verbal intelligence is that of rapidly retrieving lexical information, such as letter names, from memory. A few years later, Sternberg identified key mental processes alleged to underlie many cognitive tasks, especially ones involving inductive reasoning. They included encoding stimuli, inferring relations between stimuli, and applying what one has learned.

Ian Deary and colleagues have sought to understand intelligence through the study of inspection time. They have found that more intelligent individuals can discriminate the lengths of the lines with less stimulus duration (inspection) times than can less intelligent individuals.

Two leaders in the field of cognitive psychology, Allen Newell and Herbert A. Simon, used computers to model intelligence. The underlying idea is that computers can, in some sense, show a kind of intelligence similar to that shown by humans. During the late 1950s and early 1960s, they worked with a computer expert, Clifford Shaw, to construct a computer model of human problem solving. Called the "General Problem Solver," it could solve a wide range of fairly structured problems such as logical proofs and mathematical word problems. In 1972, Newell and Simon proposed a general theory of problem solving, much of which was implemented on the computer. It involved heuristics such as means–ends analysis, whereby at each step of problem solving, one seeks to reduce as much as possible the distance between the solution and where one is in the problem.

More recently, Marcel Just and Patricia Carpenter showed that complex intelligence test items, such as figural matrix problems involving reasoning with geometric shapes, could be solved by a computer program at a level of accuracy comparable to that of human test takers.

The models described heretofore are serial-processing models, whereby the computers take steps in sequence. David E. Rumelhart and Jay L. McClelland proposed what they called "parallel distributed processing" models of the mind. These models postulated that many types of information processing occur at once rather than just one at a time.

1.4. Cognitive–Contextual Theories

Cognitive–contextual theories deal with the way in which cognitive processes operate in various environmental contexts. Two of the major theories of this type were proposed by Howard Gardner and Sternberg.

In 1983, Gardner proposed a theory of what he called "multiple intelligences." In the 1999 version of the theory, the multiple intelligences include (at a minimum) linguistic, logical–mathematical, spatial, musical, bodily–kinesthetic, naturalist, interpersonal, and intrapersonal intelligences. Gardner also speculated as to the existence of an existential intelligence.

An alternative theory that also takes into account both cognition and context is Sternberg's "triarchic" theory of successful intelligence. According to Sternberg, intelligence is the ability to succeed according to one's own definition of success within one's sociocultural context. People succeed by recognizing and capitalizing on strengths and by recognizing and either compensating for or correcting weaknesses. Intelligence has three

aspects. These aspects relate intelligence to what goes on internally within a person, to what goes on in the external world, and to experience that mediates between the internal and external worlds.

The first aspect is the set of cognitive processes and representations that form the core of all thought. Sternberg distinguished among three kinds of processes: those involved in deciding what to do and later in deciding how well it was done, those involved in doing what one had decided to do, and those involved in learning how to do it in the first place. The second aspect is the application of these processes to the external world. According to Sternberg, mental processes serve three functions in the everyday world: adapting to existing environments, shaping existing environments into new ones, and selecting new environments when old ones prove to be unsatisfactory. According to the theory, more intelligent persons not only can execute many cognitive processes quickly or well but also know what their strengths and weaknesses are and capitalize on the strengths while remedying or compensating for the weaknesses. More intelligent persons, then, find a niche in which they can operate most efficiently.

The third aspect of Sternberg's triarchic theory is the integration of the internal and external worlds through experience. One measure of intelligence is the ability to cope with relatively novel situations. The abilities to cope with relative novelty and to automatize cognitive processing are seen as interrelated; the more a person is able to automatize the tasks of daily life, the more mental resources there are left to cope with novelty.

In 1990, Peter Salovey and John Mayer proposed the construct of emotional intelligence. This construct was then popularized by Daniel Goleman in a 1995 book. Emotional intelligence is the ability to perceive accurately, appraise, and express emotion; the ability to access and/or generate feelings when they facilitate thought; the ability to understand emotion and emotional knowledge; and the ability to regulate emotions to promote emotional and intellectual growth. Several tests are now available to measure emotional intelligence. They generally show modest to moderate correlations with conventional tests of intelligence.

1.5. Biological Theories

Biological theories are based in the neuropsychological functioning that produces, at least in part, intelligent behavior. Several biological approaches have been used. One approach has been the investigation of types of intellectual performance as related to the regions of the brain from which they originate. A researcher in this area, Jerre Levy, found that the left hemisphere is superior in analytical functioning, with the use of language being a prime example. The right hemisphere is superior in many forms of visual and spatial performance, and its functioning tends to be more synthetic and holistic, compared with the left hemisphere.

A second approach to research has involved the use of brain wave recordings to study the relation between these waves and either performance on ability tests or performance in various kinds of cognitive tasks. Researchers have found a relationship between certain aspects of electroencephalogram (EEG) and event-related potential (ERP) waves and scores on standard psychometric tests of intelligence.

A third approach involves the measurement of blood flow in the brain, which is a fairly direct indicator of functional activity in brain tissue. In such studies, the amount and location of blood flow in the brain is monitored while participants perform cognitive tasks. Horn found that older adults show decreased blood flow to the brain relative to younger adults. He also found that such decreases are greater in some areas of the brain than in others and that the decreases are particularly notable in those areas that are responsible for close concentration, spontaneous alertness, and the encoding of new information. Using positron emission tomography (PET), Richard Haier discovered that people who perform better on conventional tests of intelligence often show less activation in relevant portions of the brain than do those who do not perform as well. Presumably, this pattern of results reflects the fact that the better performers find the tasks to be easier and, thus, invoke less effort than do the poorer performers.

2. DEVELOPMENT OF INTELLIGENCE

There have been diverse approaches to studying the development of intelligence. Psychometric theorists, for instance, have sought to understand how intelligence develops in terms of changes in the factors of intelligence over time and changes in the amounts of the various abilities that children have. For example, the concept of mental age was popular during the first half of the 20th century. A given mental age was believed to represent an average child's level of mental functioning for a given chronological age. Thus, an average 10-year-old would have a mental age of 10 years, but an above-average

8-year-old or a below-average 12-year-old might also have a mental age of 10 years. The concept of mental age has fallen into disfavor, however, and is now used only rarely. The concept does not work well over a chronological age of roughly 16 years, and its assumption of perfectly smooth continuous mental development is questionable.

2.1. Piaget's Theories on Intelligence

Jean Piaget, a Swiss psychologist, suggested that the child explores the world, observes regularities, and makes generalizations much as a scientist does. Two fundamental cognitive processes, assimilation and accommodation, are alleged to work in somewhat reciprocal fashion. Assimilation involves incorporating new information into an already existing cognitive structure. Accommodation involves revising or adapting a cognitive structure to incorporate new information. Cognitive development, according to Piaget, represents a dynamic equilibrium between these two processes of assimilation and accommodation.

Piaget also postulated that there are four major periods in intellectual development: sensorimotor, preoperational, concrete–operational, and formal–operational. The first, the sensorimotor period, extends from birth until roughly 2 years of age. During this period, the child learns how to modify reflexes to make them more adaptive, to coordinate actions, to retrieve hidden objects, and (eventually) to begin representing information mentally. The second period, the preoperational period, is from approximately 2 to 7 years of age. During this period, the child experiences the growth of language and mental imagery. The child also learns to focus on single perceptual dimensions such as colors and shapes. The third period, the concrete–operational period, is from approximately 7 to 12 years of age. During this period, the child develops an important set of skills that are referred to as conservation skills. The child will recognize that substances stay the same in amount, regardless of their form. Finally, children emerge into the fourth period, the formal–operational, period, which begins at approximately 12 years of age and continues throughout life. The formal–operational child develops thinking skills in all logical combinations and learns to think with abstract concepts.

2.2. Post-Piagetian Theories

Later, so-called "post-Piagetian" theories of intellectual development, which have been heavily influenced by Piaget (although they have substantially departed from his ideas), have taken several courses. One has been to expand on Piaget's work by suggesting a possible fifth (adult) period of development such as problem finding. A second course has been to suggest periods of development related to, but different from, those suggested by Piaget. A third course has been to accept that intellectual development occurs through the periods that Piaget proposed but to hold that the cognitive bases of development differ from those recognized by him. Some of these theories emphasize the importance of memory or knowledge rather than reasoning capacity.

2.3. Vygotsky's Theories on Intelligence

Lev S. Vygotsky, a Soviet psychologist, suggested that intellectual development may be largely influenced by a child's interactions with others. That is, a child sees others thinking and acting in certain ways and then internalizes and models what is seen. Vygotsky also proposed the notion of a zone of proximal development (ZPD), which is the range of ability between what a person can do on his or her own versus what the person can do with socially derived guidance. The zone is not directly observable but may be detected by providing a context in which the individual can perform both with and without guidance. The ZPD is sometimes termed the zone of potential development.

3. MEASURING INTELLIGENCE

3.1. Early Historical Background

In 1859, the publication of Charles Darwin's *The Origin of Species* had a profound effect on many lines of scientific work. Darwin suggested that the capabilities of humans are, in some sense, continuous with those of lower animals. Hence, they can be understood through scientific investigation. One person who was strongly influenced by Darwin's thinking was Sir Francis Galton, Darwin's cousin. For 7 years, between 1884 and 1890, Galton maintained a laboratory at the South Kensington Museum in London. For a small fee, visitors could have themselves measured on a variety of psychophysical tasks. These tasks included weight discrimination and sensitivity to musical pitch. Galton believed that these tests measured more than just psychophysical abilities. He believed that psychophysical abilities are the basis of intelligence and, hence,

that his tasks were measures of intelligence. Therefore, Galton's intelligence test required a person to perform simple tasks such as deciding which of two weights was heavier and showing how forcefully the person could squeeze his or her hand. The Galtonian tradition was taken to the United States by the psychologist James McKeen Cattell.

3.2. The IQ Test

A more influential tradition of mental testing was developed in France by Alfred Binet and his collaborator, Theodore Simon. In 1904, the minister of public instruction in Paris named a commission to study or create tests that would ensure that mentally retarded children received an adequate education. The minister was also concerned that certain children were being placed in classes for the retarded not because they were retarded but because they had behavioral problems and teachers did not want them in their classrooms. Binet and Simon proposed that tests of intelligence should measure skills such as judgment, comprehension, and reasoning—the same kinds of skills measured on most intelligence tests today. Binet's early test was taken to the United States by a Stanford University psychologist, Terman, whose version came to be called the Stanford–Binet test. This test has been revised frequently and continues in use.

The Stanford–Binet test and others like it have traditionally yielded, at the very least, an overall score referred to as an intelligence quotient (IQ). In its most recent form, this test yields an overall IQ and other scores as well, as does the Wechsler Adult Intelligence Scale (WAIS) and the Wechsler Intelligence Scale for Children (WISC).

More recent tests of intelligence have expanded the range of abilities tested. For example, in 1997, J. P. Das and Jack Naglieri produced a test, the Cognitive Assessment System, based on a theory of intelligence first proposed by a Russian psychologist, Alexander Luria. The test measures planning abilities, attentional abilities, and simultaneous and successive processing abilities.

IQ was originally computed as the ratio of mental age to chronological (physical) age multiplied by 100. Thus, it is sometimes referred to as a ratio IQ. For example, if an 8-year-old child had a mental age of 10 years (i.e., performed on the test at the level of an average 10-year-old), the child would be assigned an IQ of (10/8) $\times$ 100 or 125. If the 8-year-old had a mental age of 6 years, the child's IQ would be (6/8) $\times$ 100 or 75. A score of 100, whereby the mental age equals the chronological age, would be average.

As discussed previously, the concept of mental age has fallen into disrepute, and few tests continue to involve the computation of mental ages. Many tests still yield an IQ, but it is most often computed on the basis of statistical distributions. The scores are assigned on the basis of what percentage of people in a given group would be expected to have a certain IQ. Scores computed in this way are called deviation IQs.

3.3. The Distribution of IQ Scores

Intelligence test scores follow an approximately normal distribution, meaning that most people score near the middle of the distribution of scores. Scores drop off fairly rapidly in frequency in either direction from the center of the distribution. For example, on the IQ scale, approximately two-thirds of all scores fall between IQs of 85 and 115 and approximately 95% of all scores fall between IQs of 70 and 130. Put another way, only 1 of 20 scores differs from the average IQ (100) by more than 30 points.

It has been common to associate certain levels of IQ with labels. For example, at the upper end, the label "gifted" is sometimes assigned to people with IQs over a certain point such as 130. At the lower end, mental retardation has been classified into various degrees depending on IQ so that, for example, IQs of 70 to 84 have been classified as borderline retarded, IQs of 55 to 69 as mildly retarded, IQs of 40 to 54 as moderately retarded, IQs of 25 to 39 as severely retarded, and IQs below 25 as profoundly retarded.

Many psychologists now believe that IQ represents only a part of intelligence and that intelligence is only one factor in both mental retardation and giftedness. Most current definitions of mental retardation stress adaptive skills as well as IQ, and they also emphasize attributes such as creativity, motivation, and achievement in conceptions of giftedness.

4. THE HERITABILITY AND MALLEABILITY OF INTELLIGENCE

Historically, intelligence has been viewed as a more or less fixed trait. This view conceives of intelligence as something with which people are born. Thus, the function of development is to allow this genetic endowment to express itself. A number of investigators have suggested that intelligence is highly heritable and that it is transmitted through the genes. Heritability is here

defined as the proportion of individual differences variance that is genetically transmitted. Other investigators believe that intelligence is only minimally heritable, if at all. Most authorities take an intermediate position. However, all agree that heritability operates within a reaction range, meaning that a given genotype for intelligence can result in a wide range of phenotypes, depending on how environmental factors interact with genetic ones.

Several methods are used to assess the heritability of intelligence. Perhaps the most well known is the study of identical twins reared apart. For a variety of reasons, identical twins are occasionally separated at or near birth. If the twins are raised apart, and if it is assumed that when the twins are separated they are randomly distributed across environments (often a dubious assumption), the twins would heritability have in common all of their genes but none of their environment except for chance environmental overlap. As a result, the correlation between their performances on tests of intelligence can provide an estimate of the proportion of variation in test scores due to heredity. Another method of computing the hereditary effect on intelligence involves comparing the relationship between intelligence test scores of identical twins and those of fraternal twins.

It appears that roughly half of the variation in intelligence test scores is caused by hereditary influences. However, Robert Plomin and others have shown that the heritability of intelligence increases with age, suggesting that genetic factors become more important and that environmental factors become less important to individual differences in intelligence with increasing age. During adulthood, heritability may reach as high as 70% or more. The estimates are computed, for the most part, on the basis of intelligence test scores, so that the estimates are only for that part of intelligence measured by the tests.

Whatever the heritability factor of IQ, an entirely separate issue is whether intelligence can be increased. Work by a New Zealand researcher, James Flynn, has shown that scores on intelligence tests rose rather steadily throughout the world during the mid- and late 20th century. The precise reason for the increase is unknown, although speculations include better education, better nutrition, and advances in technology.

Despite the general increase in scores, average IQs continue to vary both across countries and across socioeconomic groups. For example, many researchers have found a positive correlation between socioeconomic status and IQ, although they disagree over the reason for the relationship. Most psychologists agree that differences in educational opportunities play an important role, and some investigators believe that there is a hereditary basis for the difference as well. But there is simply no broad consensus on the issue of why the differences exist. Again, the differences are based on IQ, not broadly defined intelligence.

No matter how heritable intelligence is, some aspects of it are still malleable. Heritability of a trait is a separate issue from malleability of the trait. For example, height is highly heritable but also highly modifiable by nutrition and other environmental factors. Thus, with intervention, even a highly heritable trait can be modified. There is a growing body of evidence that aspects of intelligence can be modified as well. In the view of many psychologists, intelligence is not merely a fixed trait. A program of training in intellectual skills can increase some aspects of a person's level of intelligence. No training program—no environmental condition of any sort—is likely to make a genius of a person with low measured intelligence. But some gains are possible, and a number of programs have been developed for increasing intellectual skills. A main trend for psychologists working in the intelligence field has been to combine testing and training functions so as to enable people to optimize their intelligence.

See Also the Following Articles

Intelligence and Culture ■ Intelligence Assessment ■ Intelligence, Emotional ■ Psychometric Tests

Further Reading

Binet, A., & Simon, T. (1916). *The development of intelligence in children*. Baltimore, MD: Williams & Wilkins (Originally published in 1905).

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.

Ceci, S. J. (1996). *On intelligence*. Cambridge, MA: Harvard University Press.

Gardner, H. (1999). *Intelligence reframed: Multiple intelligences for the 21st century*. New York: Basic Books.

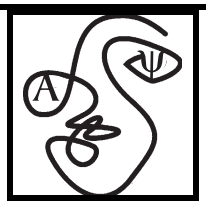
Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger/Greenwood.

Neisser, U. (Ed.). (1998). *The rising curve*. Washington, DC: American Psychological Association.

Piaget, J. (1972). *The psychology of intelligence*. Totowa, NJ: Littlefield–Adams.

Spearman, C. E. (1927). *The abilities of man*. New York: Macmillan.

Sternberg, R. J. (1990). *Metaphors of mind: Conceptions of the nature of intelligence*. New York: Cambridge University Press.
Sternberg, R. J. (1997). *Successful intelligence*. New York: Plume.
Sternberg, R. J. (Ed.). (2000). *Handbook of intelligence*. New York: Cambridge University Press.
Thurstone, L. L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.
Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.



Intentional Behavior

Martin Fishbein

Annenberg School for Communication, University of Pennsylvania, Philadelphia, Pennsylvania, USA

1. Introduction
2. Defining Intentions
3. Intentions as Predictors of Behavior
4. Moderators of the Intention–Behavior Relation
5. An Integrated Theoretical Model
6. Determinants of Behavior
7. Determinants of Intention
8. Determinants of Attitudes, Norms, and Self-Efficacy
9. The Role of Distal Variables
10. Conclusion
    Further Reading

GLOSSARY

attitude One's overall feeling of favorableness or unfavorableness with respect to a given object or action.

behavior A directly observable action performed with respect to some target, in some context, at some point in time.

behavioral categories Broad categories of behavior (e.g., prejudice, exercising, dieting) that are inferred from the observation of one or more behaviors.

behavioral intention A readiness to engage in a particular behavior.

behavioral norm One's perception of whether important others are (or are not) performing a given behavior.

correspondence The principle that measures of behavioral determinants and behavior involve exactly the same action, target, context, and time elements.

goals Desired conditions, occurrences, or states that may be outcomes of performing one or more behaviors (e.g., losing weight, staying healthy, getting an A on an exam).

meta-analysis A statistical procedure to quantify and summarize results from a number of different studies.

self-efficacy One's perception that he or she can perform a given behavior, even in the face of various obstacles.

subjective norm One's perception of what important others think he or she should (or should not) do.

Although it has been argued that many behaviors are performed automatically, without reflection on or awareness of intention, it would be hard to argue that most behaviors of interest to an applied psychologist are either automatic or habitual. In fact, most people can quickly and easily report the likelihood that they will or will not perform given behaviors, and when these measures of intention (or self-prediction) are assessed appropriately, they are highly related to actual behavioral performance.

1. DEFINING BEHAVIOR

Although behavior is often taken as a given, it is important to distinguish among behaviors, behavioral categories, and goals. All too often, investigators fail to distinguish between behaviors per se and occurrences that may be the outcome of those behaviors. For example, weight loss and success (or failure) on an exam have often been used as behavioral criteria. Unfortunately, neither of these is a behavior. Success on an exam is a possible outcome of specific actions

©2004 Elsevier Inc.
All rights reserved.

such as attending lectures, reading books, memorizing materials, and even copying answers from another person's test paper. Another problem is that investigators often treat inferences from behaviors as if they themselves were behaviors. For example, aggression, discrimination, dieting, and exercising have often been viewed as behavioral criteria. Unfortunately, none of these is a directly observable behavior. Instead, these are broad behavioral categories that are inferred from the observation of one or more behaviors. That is, although it is possible to observe specific actions that are assumed to be instances of a general class or category of behavior, the category per se cannot be directly observed. Thus, knowing that someone was "exercising," for example, might reveal very little about the specific "exercise" behaviors that the person did (or did not) perform.

In contrast to goals (or outcomes) and behavioral categories, behaviors are directly observable actions. Note, however, that the direct observation of any behavior always occurs in a given context at a given point in time. This implies that both context and time should be included in the definition of a behavior. Moreover, because most actions are directed at some object or target, a behavioral definition should involve four elements: the action (e.g., enlisting, buying, using), the target (e.g., the army, cigarettes, condoms), the context (e.g., after graduating high school, at the supermarket, for vaginal sex with one's spouse), and the time (e.g., yesterday, during the past month). Clearly, a change in any one of the elements changes the behavior under consideration. For example, enlisting in the army is a different behavior from enlisting in the navy (a change in target), using a condom for vaginal sex with one's spouse is a different behavior from using a condom for vaginal sex with a new partner (a change in context), and asking someone whether he or she purchased a car during the past 3 months is assessing a different behavior from asking whether someone purchased a car during the past 2 years.

Although all four elements should be considered in arriving at one's behavioral criterion, the level of specificity or generality used to define each element should be determined by the substantive questions being asked. Thus, a behavior may be assessed at a fairly general level (e.g., people could be asked whether they have ever [time] purchased [action] cigarettes [targets] [context unspecified]) or at a more specific level (e.g., people could be asked whether they had purchased [action] Marlboro Light 100s [target] from their local supermarket [context] during the past 2 weeks [time]).

2. DEFINING INTENTIONS

Intentions can be conceptualized as a readiness to engage in a particular behavior. This readiness to act can be assessed by asking people to indicate the extent to which they agree (or disagree) with statements such as "I will engage in the behavior," "I intend to engage in the behavior," "I expect to engage in the behavior," "I am willing to engage in the behavior," and "I will try to engage in the behavior." A number of investigators, however, have proposed that some of these items assess distinct constructs. For example, it has been suggested that it is important to distinguish between intention and behavioral expectation. More specifically, it was hypothesized that behavioral expectations are better predictors of behavior than are behavioral intentions because the former are more likely to take into account possible impediments to performance of the behavior. Thus, items such as "I intend to . . .," "I will try to . . .," and "I plan to . . ." have been used to assess intentions and items such as "I expect to . . ." and "I will . . ." to assess behavioral expectations. Available evidence to date suggests that there is little to be gained by this proposed distinction. For example, a meta-analysis of a broad set of behaviors found no difference in the predictive validity of expectations and intentions. Thus, this article uses the term "behavioral intention" to refer to measures that assess one's readiness to engage in a specific behavior. In addition, it should be noted that people may also form intentions to engage in a behavioral category (e.g., to exercise) or to pursue some goal (e.g., to lose weight).

3. INTENTIONS AS PREDICTORS OF BEHAVIOR

Empirical research over the past three decades or so has led to the recognition that behaviors can be predicted with considerable accuracy by appropriately assessing intentions to engage in the behaviors under consideration. For a measure of intention to correlate closely with behavior, the intention measure must involve exactly the same elements as does the behavior itself.

3.1. The Principle of Correspondence

The requirement that measures of intention and behavior involve exactly the same action, target, context, and time elements is known as the principle of correspondence or compatibility. The principle of correspondence is sometimes misunderstood to mean that intentions and behaviors should always be defined and measured at a very specific level. In reality, the principle simply suggests that the two variables—intention and behavior—should be measured at equivalent levels of generality or specificity. If, for whatever reason, investigators are interested in a very specific behavior performed in a particular context and at a given point in time, the intention should be assessed at that same level of specificity. However, investigators might not be interested in a particular context or time, and they can generalize across these elements in their measures of behavior and in their measures of intention. To the extent that the assessments of intention and behavior comply with the principle of correspondence, they should correlate highly with each other.

Many studies have substantiated the predictive validity of behavioral intentions. When measured appropriately, behavioral intentions account for an appreciable proportion of variance in actual behavior. Meta-analyses covering diverse behavioral domains have reported mean intention–behavior correlations ranging from .45 to .62. Studies in specific behavioral domains, such as condom use and exercise, have produced similar results, with intention–behavior correlations ranging from .44 to .56. A meta-analysis of these and other meta-analyses reported an overall correlation of .53 between intention and behavior.

Thus, there is considerable evidence that the performance or nonperformance of a given behavior is determined primarily by the strength of a person's intention to perform (or to not perform) that behavior, where intention is defined as the subjective likelihood that one will perform (or try to perform) the behavior in question. In contrast to the strong relationship between behavioral intention and behavior, however, intentions to reach goals and intentions to engage in behavioral categories might not provide very good prediction of goal attainment or the performance of a specific behavior within the behavioral category. In the latter case, the problem is largely methodological and in many ways is nothing more than a case of failure to obtain correspondence. That is, people are asked their intentions to engage in a class of behaviors (e.g., to diet or exercise), but the behavioral criterion is a measure of whether they have or have not performed one or more specific behaviors that are assumed to be a good indicant of the category. With respect to goal attainment, it should be clear that achieving a goal often requires more than a behavioral performance on the part of an individual. Achieving a goal often depends on the decisions or actions of other people or on the occurrence of certain events. To put this somewhat differently, goal (or outcome) attainment is often not completely under an individual's control, and the less control one has over attaining some goal, the lower the intention–goal attainment relation. In the same sense, one could argue that the less a given behavior is under one's control, the lower the intention–behavior relationship. Fortunately, as pointed out earlier, most human behaviors (and certainly most that are of interest to an applied psychologist) are under volitional control. Nevertheless, people do not always act in accordance with their intentions.

4. MODERATORS OF THE INTENTION–BEHAVIOR RELATION

Perhaps the two major factors that prevent one from acting on his or her intentions are (a) a lack of necessary skills and abilities and (b) the presence of environmental constraints (or the absence of environmental facilitators). Other factors may also account for observed low intention–behavior relations. As discussed previously, low correlations may be due to a lack of correspondence between the measures of intention and behavior. In addition, intentions may change over time; the longer the time interval between the assessment of intention and the assessment of behavior, the greater the likelihood that the intention will change. This implies that the longer the interval between the assessment of intention and the observation of behavior, the lower the intention–behavior correlation.

5. AN INTEGRATED THEORETICAL MODEL

Although there are many theories of behavioral prediction and behavior change in the literature (e.g., the theory of planned behavior, the theory of subjective culture and interpersonal relations, the transtheoretical

model of behavior change, the information/motivation/behavioral skills model, the health belief model, social cognitive theory, the theory of reasoned action), a careful consideration of these theories suggests that there are only a limited number of variables that need to be considered in predicting and understanding any given behavior.

5.1. Attitudes, Perceived Norms, and Self-Efficacy

In general, one can identify three factors that may directly influence an individual's intentions and behaviors: (a) the individual's attitude toward performing the behavior (i.e., the person's overall positive or negative feelings of favorableness or unfavorableness with respect to performing the behavior) that is based on the person's beliefs that performing the behavior will lead to various positive or negative consequences (or outcomes), (b) perceived normative pressure that includes the perception that the individual's important others think that one should (or should not) perform the behavior in question as well as the perception that these important others are (or are not) themselves performing the behavior, and (c) self-efficacy or personal agency that involves the individual's perception that he or she can perform the behavior in question under a variety of difficult or challenging circumstances. These and other considerations suggest the following integrative model of behavioral prediction (Fig. 1).

6. DETERMINANTS OF BEHAVIOR

In Fig. 1, it can be seen that any given behavior is most likely to occur if one has a strong intention to perform the behavior, if one has the necessary skills and abilities required to perform the behavior, and if there are no environmental constraints preventing behavioral performance. Indeed, if one has made a strong commitment (or has formed a strong intention) to perform a given behavior, if one has the necessary skills and abilities to perform the behavior, and if there are no environmental constraints to prevent the performance of that behavior, the probability is very high that the behavior will be performed.

In some populations or cultures, the behavior might not be performed because people have not yet formed intentions to perform the behavior. In other populations,

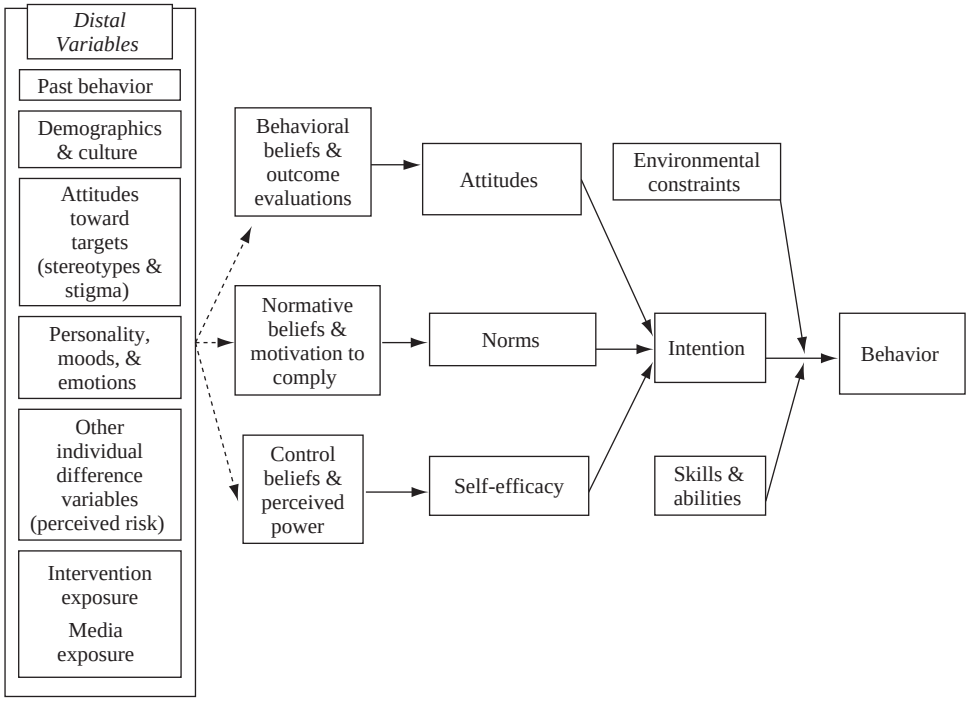


FIGURE 1 An integrative model of behavioral prediction.

the problem might be a lack of skills and/or the presence of environmental constraints. In still other populations, more than one of these factors may be relevant. Clearly, if people have formed intentions but are not acting on them, behavioral performance will depend on enhancing their skills and abilities or on removing (or learning to overcome) environmental constraints.

7. DETERMINANTS OF INTENTION

On the other hand, if strong intentions to perform the behavior in question have not been formed, the model suggests that there are three primary determinants of intention: the attitude toward performing the behavior, perceived norms concerning performance of the behavior, and one's self-efficacy with respect to performing the behavior. It is important to recognize that the relative importance of these three psychosocial variables as determinants of intention will also depend on both the behavior and the population being considered. For example, one behavior may be determined primarily by attitudinal considerations, whereas another behavior may be influenced primarily by feelings of self-efficacy. Similarly, a behavior that is attitudinally driven in one population or culture may be normatively driven in another population or culture. Thus, to understand intentional behavior, it is important to first determine the degree to which that intention is under attitudinal, normative, or self-efficacy control in the population in question.

8. DETERMINANTS OF ATTITUDES, NORMS, AND SELF-EFFICACY

The model in Fig. 1 also recognizes that attitudes, perceived norms, and self-efficacy are themselves functions of underlying beliefs—about the outcomes of performing the behavior in question, about the normative proscriptions and/or behaviors of specific referents, and about specific barriers to behavioral performance.Thus, the more one believes that performing the behavior in question will lead to "good" outcomes and prevent "bad" outcomes, the more favorable one's attitude toward performing the behavior should be. Similarly, the more one believes that specific others are performing the behavior and/or the more these others think that one should (or should not) perform the behavior in question, the more social pressure one will feel with respect to performing (or not performing) the behavior. Finally, the more one perceives that he or she can (i.e., has the necessary skills and abilities to) perform the behavior even in the face of specific barriers or obstacles, the stronger one's self-efficacy will be with respect to performing the behavior.

It is at this level of underlying beliefs that the substantive uniqueness of each behavior comes into play. Clearly, a change in any one element in the behavioral definition will usually lead to very different beliefs about the consequences of performing that behavior, about the expectations of relevant others, and about the barriers that may impede behavioral performance and, thus, will usually lead to very different attitudes, subjective norms, and perceptions of self-efficacy and intentions. For example, the barriers to, and the outcomes (or consequences) of, always using a condom for vaginal sex with an occasional partner may be very different from those associated with always using a condom for vaginal sex with one's spouse. Yet it is these specific beliefs that must be considered if one wishes to fully understand intentions and behavior. Although an investigator can sit in her or his office and develop measures of attitudes, perceived norms, and self-efficacy, the investigator cannot know what a given population (or a given person) believes about performing a given behavior. Thus, one must go to members of that population to identify salient outcome, normative, and efficacy beliefs. To put this somewhat differently, one must understand the behavior from the perspective of the population that one is considering.

9. THE ROLE OF DISTAL VARIABLES

Finally, Fig. 1 also shows the role played by more traditional demographic, economic, personality, attitudinal, and other individual difference variables (e.g., perceived risk, sensation seeking). Although it is clear that these types of variables may often be important determinants of a given behavior, the model suggests that these types of variables play primarily an indirect role in influencing behavior. That is, these "background" factors may or may not influence the behavioral, normative, or self-efficacy beliefs underlying attitudes, norms, or self-efficacy. For example, although men and women may hold different beliefs about performing some behaviors, they may hold very similar beliefs with respect to performing other behaviors. Similarly, rich and poor, old and young, those from developing countries and those from developed countries, those who plan to go to college and those

who do not, those with favorable attitudes toward law enforcement and those with unfavorable attitudes toward law enforcement, those who are happy and those who are sad or angry, those who have used drugs and those who have not—all may hold different attitudinal, normative, or self-efficacy beliefs with respect to one behavior but may hold similar beliefs with respect to another behavior. In other words, there is no necessary relation between these "distal" or "background" variables and any given behavior. Nevertheless, distal variables, such as cultural and personality differences, moods and emotions, and differences in a wide range of values, should be reflected in the underlying belief structure.

10. CONCLUSION

In general, most human behavior is intentional; that is, people typically perform behaviors that they intend (or plan) to perform, and they do not perform behaviors they do not intend to perform. In addition, they usually intend to perform (or not perform) a behavior for one or more of three reasons: they think that performing the behavior is a "good" thing to do, they feel strong social pressure to perform the behavior, or they believe that they have the necessary skills and abilities to perform the behavior. These attitudes, perceived norms (or social pressure), and feelings of efficacy are themselves determined by underlying beliefs—about the expected outcomes of performing the behavior, about the expectations and behaviors of others, and about their ability to overcome specific barriers that could impede behavioral performance. The full understanding of the performance or nonperformance of a given behavior ultimately rests on understanding which of these underlying beliefs is (are) responsible for the performance or nonperformance of that behavior.

See Also the Following Articles

Self-Control

Further Reading

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior & Human Decision Processes*, *50*, 179–211.

Ajzen, I., & Fishbein, M. (1977). Attitude–behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin*, *84*, 888–918.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Becker, M. H. (1974). The health belief model and personal health behavior. *Health Education Monographs*, *2*, 324–508.

Fishbein, M. (2000). The role of theory in HIV prevention. *AIDS Care*, *12*, 273–278.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading, MA: Addison–Wesley.

Fishbein, M., Triandis, H. C., Kanfer, F. H., Becker, M., Middlestadt, S. E., & Eichler, A. (2001). Factors influencing behavior and behavior change. In A. Baum, & T. A. Revenson (Eds.), *Handbook of health psychology*. Mahwah, NJ: Lawrence Erlbaum.

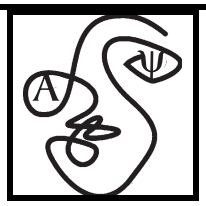
Fisher, J. D., & Fisher, W. A. (1992). Changing AIDS-risk behavior. *Psychological Bulletin*, *111*, 455–474.

Petraitis, J., Flay, B. R., & Miller, T. Q. (1995). Reviewing theories of adolescent substance use: Organizing pieces in the puzzle. *Psychological Bulletin*, *117*, 67–86.

Prochaska, J. O., & DiClemente, C. C. (1992). Stages of change in the modification of problem behaviors. *Progress in Behavior Modification*, *28*, 184–218.

Sheeran, P. (2002). Intention–behavior relations: A conceptual and empirical review. In W. Stroebe, & M. Hewstone (Eds.), *European review of social psychology* (Vol. 12, pp. 1–36). Chichester, UK: Wiley.

Triandis, H. C. (1972). *The analysis of subjective culture*. New York: John Wiley.



Intergroup Relations and Culture

Karmela Liebkind
University of Helsinki, Helsinki, Finland

1. In the Crossroads of Various Research Traditions
2. Culture: The Missing Link in Intergroup Relations
3. Intercultural Contact as Interpersonal Adjustment
4. The Social Psychology of Intergroup Bias
5. Culture as a Moderator of Intergroup Bias
6. The Role of History and Social Context in Intergroup Bias
7. The Reduction of Intergroup Bias
   Further Reading

GLOSSARY

contact hypothesis A theoretical framework that assumes that intergroup relations can, under certain conditions, be improved by increasing intergroup contact.
cultural distance Cultural differences between two groups; the more different the languages, family structures, religion, standard of living, and values of the two groups, the greater the cultural distance between them.
discrimination Unequal treatment of members of a social group on unacceptable grounds (e.g., ethnic/cultural origin, language, religion, skin color).
immigrants Persons of foreign nationality moving to a country for the purpose of staying there; in time, they may form new ethnic groups in the society of settlement.
minority groups Social groups in a numerical minority in the society; numerical inferiority is often combined with low status, but there are also high-status minorities.
prejudice Negative stereotypical beliefs about a social group combined with negative feelings toward that group.
social identity Those aspects of an individual's self-image that derive from the social categories to which the individual perceives himself or herself as belonging; these aspects contribute positively or negatively to the individual's self-esteem.
stereotyping A generalizing view of a social group that ignores individual differences among the group members.

Intergroup relations have traditionally been studied by social psychologists, whereas intercultural relations have traditionally been studied by cross-cultural psychologists. All humans can fall prey to stereotyping, discrimination, and racism, but culture determines acceptable levels of expressed bias. Collective representations about the past influence current intergroup relations. Minority members have differing attitudes about retaining their culture of origin and becoming part of the mainstream society These attitudes interact with those of the majority population and with official minority policies. The most harmonious intergroup relations are predicted when both majority and minority members share integration, assimilation, or individualist orientations. Explicit outgroup attitudes and stereotyping operate in a conscious mode, but there are also attitudes and stereotypes that function in an unconscious fashion. The translation of prejudiced beliefs (stereotypes) and evaluations (attitudes) into action (discrimination) depends on both personal and social factors. Thus, both individual and intergroup approaches can make important contributions to the reduction of intergroup bias. Interventions used for reducing bias usually focus on the need to increase the quantity and enhance the quality of intergroup contact. Some interventions can be effective

©2004 Elsevier Inc.
All rights reserved.

under particular contact conditions, whereas others may fail, particularly with respect to generalization from the individual out-group member to the out-group as a whole and to finding an intervention that works for both majority and minority groups.

1. IN THE CROSSROADS OF VARIOUS RESEARCH TRADITIONS

Issues on cultural pluralism or multiculturalism, ethnicity, and migration are hotly debated in many countries. These issues pose many political, legal, social, and economic questions. In addition, they raise many social psychological questions regarding acculturation, integration, stereotyping, "new" or "subtle" racism, ethnic identity, and intergroup relations. It is obvious that a clear understanding of intergroup relations can facilitate the processes of immigration, acculturation, and integration.

Intergroup relations have been studied mostly by social psychologists, whereas intercultural relations have been studied mostly by cross-cultural psychologists, but traditionally these two research traditions have largely tended to ignore each other. In addition, generations of researchers in the field of intergroup relations have alternated individual versus contextual levels of analysis, sometimes pushing them as far as they can go before collapsing under the accumulated complexity of evidence (pro and con) and forgetting the advantages of previous approaches. Individual difference approaches to stereotyping, prejudice, and racism all apparently emerge from U.S. researchers. Coupled with an explicit constitutional ideology of equality, the U.S. cultural focus on individualism places responsibility for bias on individuals and privileges individual autonomy over ethnic group identity. In Europe, the histories of intergroup encounters are centuries older. Coupled with the importance of linguistic, cultural, religious, and geographical boundaries, a lesser focus on individualism makes Europe the logical birthplace of more contextual approaches to intergroup relations.

Intergroup behavior can be defined as actions by members of one group toward members of another group. The most studied social psychological phenomenon related to intergroup relations is intergroup bias. Forms of intergroup bias range from prejudice and stereotyping (via discrimination, injustice, perpetuation of inequality, and oppression) to ethnic cleansing and genocide. Various forms of prejudice (i.e., the holding of derogatory attitudes about members of a social group) are more common than outright manifestations of intergroup conflict. In practice, therefore, the vast majority of social psychological studies have investigated weaker forms of bias, as expressed by persons with relatively mild prejudice.

Within both social and cross-cultural psychology, it has been customary to look at social behavior from a perspective that reflects the predominantly individualist emphasis of many contemporary North American researchers. Research on intergroup relations, however, more often applies a perspective that still focuses on individual behavior but that takes group membership as the determinant of that behavior. The individualist approach sees group behavior as an aggregated result of the social cognitions and resultant behaviors of its members. The more contextual approach reverses this presumed causal chain of events in that the individual's identity and consequent behaviors are seen as defined by his or her group memberships. The principal theorists proposing this perspective have been Europeans, but the influence of their ideas is now widely diffused. Most researchers would probably agree that social psychology should be able to offer explanations of both individual and small group behavior. European researchers tend to include explanations at the level of intergroup relations and of society as a whole. It is through an emphasis on these more macroscopic levels of influence that recent European work has become distinctive.

However, social psychology has also been criticized for being traditionally more concerned with general processes and structures than with particular contents related to historical, cultural, and contextual circumstances. Consequently, the growing interest within social psychology in content, culture, and context is very welcome. Perceptions of history and common ancestry, as well as assumed cultural characteristics, play a central role in interethnic/intercultural relations. Many questions and problems in studying cultural features are discussed extensively in cross-cultural and cultural psychology. It is the contextual, rather than the individualist, approach to intergroup and intercultural relations that dominates the bulk of this article.

2. CULTURE: THE MISSING LINK IN INTERGROUP RELATIONS

2.1. What Is Culture?

Various aspects of the environment are more or less important for members of different cultures. Culture

has been defined in numerous ways. It can be seen as consisting of a set of human-made objective and subjective elements that in the past have increased the probability of survival and resulted in satisfaction for participants in an ecological niche and, thus, have become shared among those who could communicate with each other because they had a common language and they lived in the same time and place. Cultural differences can also be seen as being tied to historical experiences, some of which are based on socioeconomic stratification of a relatively recent nature and others of which are rooted in a distant path of which nobody is conscious. This historically older culture is sometimes called "deep culture" because it is embedded in language, ethnicity, religion, and/or nationality as such.

For the purpose of this article, culture is conceptualized as shared patterns of belief and feeling toward issues such as child-rearing practices, family systems, and ethical values or attitudes. These are elements of what Triandis has called "subjective culture" that is created in interactions among people. In this perspective, there are as many cultures as there are languages and ways of life within each language (e.g., physicians have a culture that is somewhat different from that of lawyers). Thus, culture can be viewed as a relatively stable and characteristic system of shared meanings that define or determine perceptions and behavior. This approach is typically taken when different ethnic groups are compared, especially within cross-cultural psychology. Changes are not denied, but a more long-term perspective is taken. More gradual changes are examined in research on acculturation that focuses on psychological and behavioral changes that an individual or group experiences as a result of sustained contact with members of another cultural group(s).

2.2. Culture in the Research on Intergroup Bias

The groups involved in intercultural encounters are mostly large-scale ethnic or cultural groups. Yet much of even the more contextual branch of social psychological research on intergroup relations concentrates mainly on status differences between groups. Although an exclusive social position perspective offers important insights into intergroup relations, there are limits to using such a perspective. History and culture are the two main ingredients of ethnicity. Ethnicity, in turn, is most often thought of as culture that is transmitted across generations.

All humans can fall prey to stereotyping, discrimination, and racism. At a more global level, however, little is known about individual differences in these phenomena. In some Asian cultures, which have a history of being more ethnically homogenous within nations and more collective in general, social psychologists rarely study individual variation in stereotyping and prejudice. Likewise, Latin American and African social psychologists still have much to say on this topic. However, the importance of cultural and local norms in predicting discriminatory behavior should not be underestimated. Culture channels stereotyping and prejudice by defining who constitutes "us" and who constitutes "them." Culture also determines acceptable levels of expressed bias, from subtle to overt. In addition, cultures differ in norms for describing perceived differences between social categories as either inherent and traditional or unacceptable and controllable.

It must be acknowledged that the vast majority of research on intergroup issues, such as stereotyping, prejudice, and discrimination, is carried out in societies that are low on cultural measures of hierarchy, power distance, and conservatism. Such systems tolerate considerable public discussion about inequalities and their legitimacy. Given the egalitarianism of such societies—or at least their ideal of egalitarianism—one must be cautious in generalizing globally some of the psychological processes that have been identified by this research.

3. INTERCULTURAL CONTACT AS INTERPERSONAL ADJUSTMENT

3.1. Intercultural Relations and Sociocultural Adaptation

When groups of individuals having different cultures come into continuous firsthand contact with each other, the consequence is likely to be subsequent changes in the original cultural pattern of either group or both groups. This is how the process of so-called acculturation is traditionally defined. This process refers to groups in contact, thereby explicitly placing acculturation within the realm of intergroup relations. However, within cross-cultural psychology, theoretical frameworks in relation to the acculturation process have not been borrowed from social psychology; instead, they have been borrowed from different areas of mainstream psychology, notably the stress and coping literature, on the one hand, and research on social learning and skills acquisition, on the other.

In both perspectives, acculturation is nearly exclusively seen as an individual-level process of adjustment and adaptation to a new culture. As a consequence, the acculturation literature has dealt more with interpersonal relations than with intergroup relations.

Ward and colleagues have argued that adaptive outcomes of acculturation are meaningfully divided into psychological (emotional/affective) and sociocultural (behavioral) domains. The former has traditionally been the focus of study for the stress and coping framework, and the latter has traditionally been the focus of study for the social learning framework. Individuals who are exposed to acculturative demands are motivated to maintain psychological well-being as well as to acquire culturally appropriate knowledge and skills. Empirically, these dimensions are moderately and positively related. This interrelation is understandable given that successful social adaptation to a new culture arises in part from psychological well-being and leads to a feeling of greater self-efficacy. This, in turn, enhances psychological well-being and vice versa.

It is obvious that good social relations with members of the other cultural group are a vital part of sociocultural adaptation. The more different the languages, family structures, religion, standard of living, and values of the two groups, the greater the cultural distance between the groups. The greater the cultural distance, the more the norms are likely to be different and the more difficult it is for the two groups to develop good intergroup relations. Cultural distance results in perceived dissimilarity, which can lead to conflict if there are different goals and interests. Individuals confronting new environments with greater cultural dissimilarities in values will experience more difficulties with adaptation. A robust relationship has been consistently observed between cultural distance and sociocultural adjustment problems—the greater the cultural similarity, the less difficulties with adaptation. This outcome also holds for perceived degrees of cultural differences.

Scholars approaching the problem of cross-cultural adaptation from a skills perspective emphasize the role of culture-specific knowledge for sociocultural adaptation. Depending on which particular cultural groups are involved, people will need specific knowledge of the rules of interaction with their hosts. Knowledge about normative patterns of behavior in the new culture will then have to be put into practice. For example, in training programs, Arab nonverbal skills can be taught to normally "stand-offish" English persons, thereby enhancing the impressions they give to Arabs. Various types of cross-cultural training can be very effective, although some persons will already have culturally fitting skills through personalities or previous cross-cultural experiences. Powerful effects of cross-cultural training have been demonstrated for several factors such as better interpersonal skills with members of the other cultural group and better understanding of the social system and values of the other cultural group.

One important skill worth identifying separately is language competence. Greater fluency in the language of the other cultural group greatly increases the likelihood of successful intercultural communication. Of course, because the vast majority of cultural sojourners or immigrants will have no access to any cross-cultural training programs, they will have to adapt as best as they can. It is here that broad personality dispositions and the intercultural attitudes of the individuals, on the one hand, and the specific social, cultural, and historical context of the encounter, on the other, will influence the ultimate outcome of the intercultural contact.

4. THE SOCIAL PSYCHOLOGY OF INTERGROUP BIAS

4.1. Individual- and Group-Level Explanations

A popular view of the origins of prejudice is that that it is primarily a personality problem. The most influential example of this type of theory was that proposed by Adorno and colleagues in 1950. Working from a Freudian perspective, they believed that a child's natural aggression toward his or her parents (an inevitable consequence of being subjected to constraints) was displaced onto alternative targets due to the feared consequences of displaying it directly. The likely choice of targets would be those seen as weaker than or inferior to the child such as members of deviant groups or ethnic minorities. The end result was the so-called authoritarian personality. The association between authoritarianism and various forms of prejudice has been confirmed in several intergroup contexts. Recent work points to a positive relationship between prejudice and right-wing authoritarianism (i.e., submitting to established authorities and adhering to social conventions).

A similar individual-level theory of prejudice is the so-called frustration–aggression theory, which implies that aggression is often not directed at the true source of the frustration people feel (e.g., the capitalist system that causes an economic recession); instead, it is often diverted onto vulnerable and easily accessible targets

such as members of minority groups. Attempts to confirm this so-called scapegoat theory of prejudice have, however, met with mixed success. It seems that absolute levels of hardship and frustration are less potent instigators of aggression than is a sense of relative deprivation, that is, a perceived discrepancy between what one has and what one feels entitled to have.

However, it is important to note the limitations of this kind of individual difference perspective. It underestimates the importance of current social situations shaping people's attitudes and tends to neglect sociocultural determinants of prejudice. For example, Pettigrew found during the 1950s that although White South Africans showed very high levels of anti-Black prejudice, they did not appear to have particularly high levels of authoritarianism as compared with populations in other countries. In addition, the personality approach cannot explain the widespread uniformity of prejudice in certain societies or subgroups within societies. In prewar Nazi Germany, as well as in many other places since then, consistently racist attitudes and behavior were shown by hundreds of thousands of people who must have differed on most other psychological characteristics. Yet another problem concerns the historical specificity of prejudice. For example, increases in racist attacks on various ethnic minority populations in Western Europe during recent decades have sometimes taken place over the space of just a few years—much too short a time for whole generations to have adopted new forms of child-rearing practices giving rise to authoritarian and prejudiced children.

4.2. Social Identity Theory

The individual difference approaches make the assumption that people's behavior in group settings is essentially similar to their behavior in all other situations. However, in group settings, people's behavior is often qualitatively different. This led Tajfel to suggest during the late 1970s that it is important to distinguish between interpersonal behavior and intergroup behavior, where the former means acting as an individual with some idiosynchratic characteristics and a unique set of personal relationships with others and the latter means acting as a group member. Any social behavior falls somewhere along a continuum defined by these two extremes. At the group level of contextual analysis of intergroup relations, the social identity theory, originally formulated by Tajfel and colleagues, has blossomed within the European context into a primary international approach to intergroup relations.

Social identity theory proposes that the social part of one's identity derives from the groups to which one belongs. By favorably comparing attributes of one's own groups with those of out-groups, this theory suggests that a person acquires both a positive sense of who he or she is and a clear understanding of how he or she should act toward in-group and out-group members. From a social psychological perspective, therefore, intergroup discrimination can be perceived as an attempt to establish a positively valued distinctiveness for one's in-group in order to achieve or maintain a positive social identity, that is, a positive collective self.

In this perspective, prejudice is perceived as stemming partly from the normal cognitive process of categorization; that is, people accentuate differences between categories and minimize differences within categories. They view categorized groups as homogenous and privilege category-confirming information in memory as well as category-confirming covariation in judgment. This means that prejudice-consistent information is attended to and rehearsed more frequently, and is remembered more accurately, than is inconsistent information.

Studies from a broad range of nations have shown that members of various ethnic groups often maintain biased systems of belief about the positive qualities of their own groups and about the negative qualities of other groups. For example, group membership determines group-serving attributions in that such membership can completely reverse the pattern of attributions made about an individual's behavior. Positive behaviors performed by members of one's own group are believed to arise from internal dispositions, whereas their negative behaviors are seen as the result of external forces. Similarly, desirable behaviors of out-group members are seen as externally caused, whereas undesirable behaviors are seen as internally caused.

However, there is a danger in overemphasizing the categorization approach to intergroup bias. It may suggest that prejudice might be inevitable, or even desirable, for cognitive adaptation. Although perceptual schemas undoubtedly simplify information, it is not clear that linguistic categories must do so. Language can be used to both categorize and particularize. The difference between prejudiced views and unprejudiced views lies in the content of these views and not the structure. It is true that people tend to interpret information in accordance with their category-based expectations. For example, if they are confronted with information about a group that disconfirms their stereotype of that group, they tend to perceive that information as nonrepresentative of the group; that is, they use the stereotype to "explain away"

deviant information. However, what is crucial here is the "category-based expectations"; whereas these expectations regarding out-groups are likely to be negative for a prejudiced individual, they may be as positive for the in-group as for out-groups for a tolerant individual. In the latter case, the person would need to explain away (i.e., particularize) the additional negative information as opposed to the additional positive information.

4.3. The Role of Perceived Threat and Unequal Position in Intergroup Bias

An in-group is a set of people who perceive each other as having something in common. There are cultural differences in the way in which in-groups are defined; that is, the bases for judging other people as in-group members can be different. There are also degrees of being in-group members; that is, people identify with in-groups in varying degrees. Typically, the more people perceive themselves as belonging together because they experience a common fate, have a common ancestry, or are similar in their gender, age, ethnicity, geographic origin, language, religion, social class, occupation, way of life, attitudes, beliefs, norms, role definitions, and/or values, the more they see each other as in-group members. Also, in-groups are dynamic constructs; that is, people change their perceptions of in-groups and out-groups according to the situation. An out-group consists of a set of people who are not members of the in-group. Members of the in-group are likely to refer to out-group members as "them" and to other in-group members as "us".

Research has identified a number of key moderators of intergroup bias. Potential moderators range from culture to education. Higher education is generally related to less intergroup bias. Culture as a moderator is discussed in more detail later. Other moderators studied include perceived threat, power or status, group size, identification, and self-esteem. Threat is a central explanatory concept in several theories of intergroup bias, although its interpretation varies greatly. Threat can be perceived in terms of the in-group's social identity, its goals and values, its position in the hierarchy, and even its existence. Threat can be realistic (e.g., intergroup competition over scarce resources) or symbolic (e.g., blocking the in-group's values, customs, and traditions). Perceived threat increases intergroup bias.

Intergroup relations among real groups tend to involve groups of unequal power. In general, members of high- and equal-power groups show more bias than do members of low-status groups, and discrimination by members of numerical minorities with high power is especially strong. Although size, status, and power tend to become confounded outside of the laboratory, these factors have also been shown to have independent effects; that is, groups in numerical minorities express more bias than do those in numerical majorities, regardless of whether the groups are real or artificial. However, when identification is experimentally induced, both majority and minority groups show bias. Members of low-status groups show more bias when status differentials are perceived as unstable and/or illegitimate and when group boundaries are perceived as impermeable. However, low-status groups may simultaneously show out-group favoritism, especially on status-relevant evaluation and when they define their inferiority as legitimate and stable.

4.4. Identification and Strategies for Escaping an Unsatisfactory Social Identity

A central idea in social identity theory is that biased intergroup comparisons are directly linked to social identification. Presumably, the more important a group is to its members, the more bias group members should show in its favor. However, in many studies, there have been rather unstable correlations between group identification and in-group bias. Hinkle and Brown suggested that the psychological processes proposed by social identity theory might not be operative in all groups. They hypothesized that this would depend on the prevailing level of individualism or collectivism in the group or group members, on the one hand, and their inclination to engage in intergroup comparisons, on the other. Group comparisons may be more important for relational groups (e.g., sports teams) and less important for autonomous groups (e.g., families). A strong link between group identification and in-group favoritism would be expected only in groups simultaneously characterized by collectivist and relational orientations. Some support has been found for this idea.

For low-status groups, the outcome of available group comparisons is often negative for their self-esteem. One reaction to this is simply to leave the group, and there are many examples of members of "inferior" groups distancing themselves physically or psychologically from their groups. Such an individualist strategy of social mobility might not always be possible, especially if the group boundaries are relatively fixed and impermeable, as is

the case with many ethnic and religious groups. In cases such as this, social identity theory suggests that a number of other avenues may be pursued. One is to limit the comparisons made to other similar or subordinate groups. Another is to sidestep the main dimensions of comparison and to either invent new dimensions or change the value of existing dimensions. These ways are expressions of social creativity. Yet another route is to directly confront the dominant group's superiority by agitating for social and economic change.

Which of these tactics will be chosen might well depend on the prevailing social climate. If it is such that no real alternatives to the status quo may be conceived, subordinate groups are unlikely to openly challenge the existing order and attempt social change. A key factor in generating social unrest among subordinate groups is a sense of relative deprivation, either in relation to their own groups in the past or (more often) in relation to the dominant group. The implications of broader social and cultural influences for intergroup bias are discussed further in the next two sections.

5. CULTURE AS A MODERATOR OF INTERGROUP BIAS

5.1. Individualist and Collectivist Cultures and Individuals

Research on culture as a moderator of intergroup bias is often conducted within an approach to cross-cultural psychology that combines individual and cultural levels of analysis and focuses on the self-concept and values. This approach has offered a perspective, the key concepts of which are individualism and collectivism (I–C). Originally, these terms were employed in reference to dimensions that differentiate cultures, describing individualist societies as those emphasizing autonomy and emotional independence and describing collectivist cultures as those showing concern for solidarity, sharing, and interdependence.

Individualist cultures, such as those found in Western societies, are defined by certain tendencies. First, people in individualist cultures see the self as independent of groups, whereas people in collectivist cultures (e.g., those found in Asia, Latin America, and most traditional societies) see the self as an aspect of one or more groups. This implies that in collectivist cultures, the goals of the in-group tend to dominate over personal goals, in-group norms are slightly more important than personal attitudes, and communal relationships tend to be more important than relationships based on personal gain. Such differences result in different definitions of the in-group and out-group and in different pressures to conform to in-group norms, with major implications for intergroup relations.

However, Triandis and Trafimov in 2001 emphasized that individuals in different cultures include in their cognitive systems both individualist and collectivist elements in different mixtures, depending on the situation. Thus, the same constructs (I–C), when studied at the individual level of analysis, are renamed idiocentrism–allocentrism (I–A), although a similar distinction was also proposed by Markus and Kitayama during the early 1990s in their description of independent and interdependent selves. Triandis has argued that the social identity of allocentrics and collectivists is more core and salient than that of idiocentrics and individualists. Consequently, competition with out-groups and ethnocentrism also are more readily apparent in the former. These contentions have yet to be tested in the realm of ethnic perceptions and relations, but researchers consider I–C and I–A to be useful constructs in the analysis of ethnocentrism in a cross-cultural context.

5.2. Empirical Evidence and Limits of Generalizability

Most commentators would agree that the hypotheses of social identity theory are more likely to be supported in collectivist cultures than in individualist ones. However, Lee and Ward during the late 1990s found only partial support for the hypothesis of stronger ethnocentrism among allocentrics. Although allocentrics had more positive in-group attitudes than did idiocentrics, no difference was found between allocentrics and idiocentrics on out-group attitudes. Thus, in-group favoritism in both Malays and Chinese was due to the ethnocentrism bias of allocentric respondents. These findings are in broad agreement with Triandis's theorizing on the heightened significance of social identity, social categorization, and social comparison for allocentrics. The evidence for allocentric ethnocentrism, however, was limited to an in-group bias. Contrary to Triandis's suggestions, there was no distinction in the out-group attitudes held by idiocentrics and allocentrics in Lee and Ward's study.

The taxonomic model proposed by Hinkle and Brown suggests that the relationship between group identification and intergroup bias would be strongest in groups that are both relational and collectivist,

whereas this link would be weak in individualist and autonomous groups. Although there is some empirical support for this model, there is little information as to whether a relational orientation is important within collectivist cultures. Triandis suggested that members of collectivist cultures are not interested in comparing themselves with out-groups. The social comparison in which allocentrics engage might not be quite the same as that envisaged by social identity theory. Allocentric group members might not be so preoccupied with their identity as group members. Instead, they may focus on their group as a whole and on threats to its well-being.

Although Triandis has taken great pains to distinguish I–C as a cultural-level label and I–A as an individual difference measure, the terms have been used interchangeably in the cross-cultural literature. In comparative studies, the constructs frequently have been operationalized through the examination of individuals from two or more cultures either a posteriori labeled or a priori selected on the basis of some overarching conceptualization of I–C. In the more sophisticated investigations, the selection, definition, and classification are based on previous empirical research, but even this approach can pose inadvertent problems, and researchers often fall prey to the "ecological fallacy," that is, the assumption that because two cultures differ on a specific dimension, individual members of those countries will reliably differ in the same predictable manner.

Lee and Ward's study among Chinese and Malays in Singapore provides a case in point. Singapore has been previously labeled and commonly classified as a collectivist culture. Triandis has argued vigorously that members of collectivist cultures are expected to display pronounced ethnocentric attitudes. However, the ethnocentric biases that Lee and Ward found in their study were limited to allocentric subjects. This highlights the necessity of the distinction between I–C and I–A as well as the awareness and acknowledgment of theory generalizability.

6. THE ROLE OF HISTORY AND SOCIAL CONTEXT IN INTERGROUP BIAS

6.1. History and Social Context: Neglected Factors in Research

Norms, beliefs, behavioral intentions, feelings, and stereotypes determine intergroup behavior. These may be both the result of the history of previous intergroup contact and the cause of future intergroup behavior. However, they may also be affected by factors other than previous intergroup contact. Sometimes it is in the interest of certain members of a group to induce other members of that group to have negative thoughts about a different group. More generally, a variety of social factors (e.g., relative power of the in-group and out-group; zero-sum conflict resulting in competition, domination, or scapegoating) affect norms (e.g., antagonism, paternalism, persecution, submission, rebellion, repression, conciliation), stereotypes (e.g., threatening, inferior, superior, oppressive, powerful, weak), feelings (e.g., accepting/rejecting), and behavioral intentions (e.g., to hurt, to derogate, to punish, to submit, ambivalence) regarding out-groups.

Collective representations about the past influence current intergroup relations, perceptions, and behavior. One should be careful not to neglect the importance of more enduring identities with long-term commitments and connections to, for example, former and future generations. A sense of enduring group membership may be related to cultural features such as a strong collectivist value orientation in which "in-group-centrism" predominates and identification is more of a nominal variable than a continuous one. The history of intergroup conflict has generally been neglected in the social psychological literature. Especially in collectivist cultures, history is often a major factor in intergroup relations (e.g., the case of the former Yugoslavia).

Groups and group members derive their identities from particular contexts. However, the notion of context is ill defined across many psychological paradigms. For example, context can be taken to refer to the particular task or activity in which people are engaged, to historical and cultural circumstances, to general immigrant conditions, and/or to actual social situations (e.g., schools, neighborhoods, workplaces). Many existing psychological models are not specific enough for the study of ethnic minority groups. An understanding of such groups would require explicit attention to negative social circumstances, such as racism and discrimination, in relation to concrete environmental influences.

6.2. The Special Circumstances of Minority Groups

Many real-life intergroup contexts consist of groups that hold either a minority or a majority position vis-à-vis each other. In real life, numerical asymmetries often, but not necessarily, covary with power or status asymmetries.

Simon and colleagues described the so-called cognitive-affective crossfire of minority members as follows. On the one hand, numerical distinctiveness makes group membership highly salient and difficult to ignore, both for minority members themselves and for others who may continually remind them of it by word and deed. In this way, group membership becomes a much more central self-aspect for minority members than for majority members. On the other hand, minority membership entails specific risks and stressful experiences that may be largely unknown to majority members. These risks and the negative affective consequences are further exacerbated when a numerical inferiority is also associated with status and/or power inferiority. In other words, there are stronger cognitive forces pushing minority members toward their group (or keeping them in it), while at the same time there may also be stronger affective forces pulling them away from it (or keeping them out of it).

Depending on the perceived opportunity structure of the social context (e.g., the permeability of group boundaries), minority members may opt for individualist strategies involving psychological dis-identification or actual exit from their group, for collectivist strategies involving assertive intergroup behavior, or for a combination of both. Majority group members can be more "mindless" than minority group members regarding group membership, at least so long as the former do not feel threatened by an assertive minority. Both minority and majority group members are likely to adopt an intergroup perspective when minority group members show assertive intergroup behavior to escape from or cope with the cognitive-affective crossfire (or worse) because such assertiveness may incite defensive intergroup behavior on the part of majority group members. Note that, on the surface, the assertive behavior of a minority and the defensive behavior of a majority may take the same form such as intergroup discrimination. From a theoretical point, however, the underlying meanings are different because such phenotypically similar behavior serves different purposes for minority and majority members.

6.3. The Special Circumstances of Immigrants

Historical circumstances, group norms, and ideologies also play a crucial role in the intergroup relations between immigrants and host populations. Immigrant groups, as well as individual immigrants, arrive in a new country with differing attitudes about retaining their culture of origin and becoming part of the new society. In the new society, these attitudes interact with the actual and perceived levels of acceptance of immigrants and with official policies toward immigration. For example, assimilationist societies expect immigrants to abandon their own cultural and linguistic distinctiveness for the sake of adopting the culture and values of the dominant group, whereas ethnist societies often define who can be citizens of the state in exclusive terms.

The particular norms and ideologies reflected in the public rhetoric on the integration of immigrants provide the specific environmental context within which hosts and immigrants develop their cognitive representations of each other. They can perceive each other as one single in-group, as two subgroups within a common in-group, as two separate groups, or as separate individuals.

The most harmonious intergroup relations are predicted when both host community members and minority group members share integration, assimilation, or individualist orientations. In contrast, negative intergroup stereotypes and discriminatory behaviors can result when immigrant group members prefer integration but host community members insist that immigrants assimilate to the host society. However, host community members who want to segregate or exclude immigrants are likely to foster the most conflictual intergroup relations with targeted immigrant groups. These host community members are likely to have negative stereotypes regarding immigrants and to discriminate against them in many domains. Nonviolent segregationists may simply think that cultures should not mix, whereas radical exclusionists are more likely to launch violent racist attacks against immigrants. Of the targeted immigrant groups, it is those with separatist attitudes that are most likely to resist or even retaliate against host community persecutions. This promising so-called interactive acculturation model, developed by Bourhis and colleagues during the late 1990s, has received some empirical support.

6.4. Is Intergroup Bias Ubiquitous?

Stereotypes are likely to be particularly clear and consensually held if one's group has a long history of dealings with an out-group. If groups share a history of conflicts, out-group stereotypes are often extreme, simple, negative, and symmetrical, with each group rating its own group members positively while denigrating the out-group members. In general, group members rate their own group members higher on beneficence (e.g., honest, kind, loyal, trustworthy)

than they do out-group members. Out-group members do the same, resulting in a "mirror image" pattern.

The critical question for intergroup harmony is the size of the gap between the in-group and out-group ratings, especially on perceived values. Stereotypes can serve the societal function of acting as ideological representations that are used to justify and legitimize existing social and power relations within society. People can hold intense stereotypes about persons from other cultural and ethnic groups, even though they have never met them. These polarized attitudes probably arise from the home, the media, and the educational curricula. However, there are studies that have not found in-group bias in stereotyping. It is obvious that one must consider the effects of prevailing social and cultural norms on the expression of intergroup bias.

6.5. The Influence of Social and Cultural Norms on the Expression of Prejudice

Theoretically, prejudice and stereotypes are related but still distinct concepts. Attitudes are evaluative responses, whereas stereotypes are specific beliefs about a group. Responses to explicit measures of intergroup bias are given consciously. Typically, explicit attitudes are assessed by self-report measures, including attribution of group traits (stereotypes), group evaluations (prejudice), and differential behavior toward in-group and out-group targets (discrimination). However, measures of these cognitive, affective, and behavioral components, respectively, are often only modestly or weakly related in most Western societies.

Explicit attitudes and stereotyping operate in a conscious mode, whereas implicit attitudes and stereotypes are evaluations and beliefs that are automatically activated by the mere presence (actual or symbolic) of the attitude objects. Implicit attitudes commonly function in an unconscious fashion and are typically assessed using response latency procedures, memory tasks, physiological measures (e.g., galvanic skin response), and indirect self-report measures. Thus, the promise of implicit measures is to assess the true extent of people's bias given pressures to conform to socially desirable or politically correct norms. When such norms prevail, individuals may be motivated to express their prejudice in indirect ways that do not threaten their nonprejudiced self-images.

The translation of prejudiced beliefs (stereotypes) and evaluations (attitudes) into action (discrimination) depends on both personal and social factors. Highly prejudiced persons feel less guilt and shame, less ambivalence, and less internal conflict about discrepancies between internal and/or social standards of egalitarianism and actual discriminatory behavior. Less prejudiced individuals are self-critical about their departures from more egalitarian standards prevailing in their cultural and social environments. There may be a greater correspondence between implicit and explicit attitudes for issues that are not socially sensitive than for those that are socially sensitive or that are associated with more prejudiced historical norms or traditional socialization.

However, there is more to discrimination than personal proclivities. Social groups may feel that they are under political and economic threat by other groups or may be fueled by ideologies of antagonism passed down from the past and sustained by racist talk common to in-group members. These social inducements to discrimination may be further unrestrained by a wider political ideology of intergroup harmony (e.g., Canada's policies of multiculturalism and human rights observance). Such a volatile mix can be exacerbated by societal factors and cultural values supporting a sense of cultural superiority.

Discriminatory behavior need not result in open intergroup hostility if the social ideology surrounding differential access to material and social resources is accepted by members of all competing groups. Alternatively, unequal group access to resources may prevail and be resented by members of groups facing discrimination. However, the consequences of challenging this order may be too destructive to contemplate, or the social basis for group mobilization may be lacking. Conflict behavior is less in societies with crosscutting ties than in societies with ties that reinforce polarization between groups and in societies where scapegoating ideologies concerning specific out-groups are firmly entrenched.

7. THE REDUCTION OF INTERGROUP BIAS

7.1. The Contact Hypothesis

Intergroup bias has psychological, social, and cultural components; hence, both individual and intergroup approaches can make important contributions to the reduction of intergroup bias. Some interventions attempt explicitly to decrease bias on an individual level by invoking guilt for holding socially undesirable attitudes. Although this approach may sometimes seem successful, it capitalizes on the good intentions of

low-prejudiced people (whose bias is surely not the main problem) and demands appropriate levels of awareness, effort, and practice over time. Other individual-level approaches use more indirect means such as inducing empathy for out-groups.

However, most of the interventions used for reducing bias employ more intergroup methods and focus on the need to increase the quantity and enhance the quality of intergroup contact. This approach derives from the so-called contact hypothesis, which was originally developed by Allport during the 1950s and is one of the most influential ideas for the reduction of intergroup prejudice. Broadly speaking, it suggests that contact between members of different groups, under the appropriate conditions, lessens intergroup prejudice and hostility.

In 1998, Pettigrew's review of recent research pointed to the increased knowledge of the mediating processes by which contact can reduce intergroup bias and anxiety. Allport held that the positive effects of intergroup contact occur only in situations marked by four key conditions: equal group status, common goals, intergroup cooperation, and the support of authorities, laws, or customs. Many other factors have since been advanced, for example, the acquaintance potential and intimacy of the contact. Some researchers emphasize equal group status coming into the situation as part of the equality requirement, whereas others find this to be less important than group status within the situation. Social and institutional support means that those in authority (e.g., the politicians implementing new legislation and the judges and police monitoring its administration) should be unambiguous in their endorsement of measures designed to promote greater contact and integration. Creating a social climate in which more tolerant norms can emerge, and sanctioning the violation of such norms, constitutes the essence of social and institutional support.

Allport's formulation continues to receive support across a variety of situations, groups, and societies. Most studies report positive contact effects, even in situations lacking key conditions. Overall, face-to-face interaction between members of distinguishable groups is importantly related to reduced prejudice, and intergroup contact relates to a wide range of prejudice measures (e.g., belief, social distance, stereotype) and affective measures. Testing the possibility of an opposite causal sequence (i.e., that prejudiced people avoid contact more than contact lowers prejudice), it has been shown that the path from contact to reduced prejudice is stronger than the other way around. Although it concentrates on school and housing situations, the contact literature ranges from Chinese students in the United States and interracial workers in South Africa to German and Turkish school children and Americans getting to know Southeast Asian immigrants.

7.2. Mechanisms by Which Contact Reduces Intergroup Bias: The Role of Category Salience

The contact hypothesis does not specify the processes that are responsible for the positive effects or how the positive effects of contact generalize beyond the immediate situation to other situations, from the specific out-group member encountered to the out-group as a whole, or to other out-groups not involved in the contact. Recent work suggests that four interrelated processes operate through contact and mediate attitude change: learning about the out-group, changing behavior, generating affective ties, and in-group reappraisal.

Although cognitive research has uncovered a host of mechanisms that limit learning material that counters our attitudes and stereotypes, the research literature suggests that positive effects are more common than the contact hypothesis or cognitive analyses predict. The basic reason is that learning about the out-group is only one of several processes involved. For example, optimal and continuous intergroup contact has the potential to produce attitude change by revising attitudes in accordance with behavior and by reducing the anxiety often experienced in intergroup encounters. Positive emotions aroused by intergroup friendship may be pivotal in reducing intergroup prejudice. It has been shown that having an out-group friend may even generalize to other out-groups besides that of the particular friend. In addition, optimal intergroup contact reduces prejudice by providing insight about the in-group as well as the out-group, leading to a less provincial view of out-groups in general.

As for the generalization of contact across situations, research has generally shown that positive effects on intergroup attitudes might need repeated optimal situations. Research supporting the salient categorization model proposed by Hewstone and Brown during the 1980s shows that stereotype change generalizes best to the intergroup level when the individuals involved are typical out-group members. This approach argues that, to protect against loss of distinctiveness for groups involved in contact, two factors are important. First, the salience of group boundaries should be maintained during contact (i.e., group memberships are made salient or the partner is typical [rather than atypical] of the out-group as a whole) to promote generalization across members of

the target group. Second, groups should recognize and value mutual superiorities and inferiorities.

Thus, there seems to be some advantage in maintaining the original group boundaries in the interaction while otherwise optimizing contact conditions. Positive contact is more likely to be associated with favorable out-group attitudes when contact takes place with a typical out-group member. However, this poses a problem. Typical members of real groups are different in many ways, but people tend to seek out people with similar interests and status. Maintaining category distinctions during contact involves the risks of reinforcing perceptions of group differences and increasing intergroup anxiety, thereby increasing bias.

One way of maintaining category salience while still diminishing anxiety associated with intergroup contact is offered by a recent development of the contact hypothesis called the extended contact effect. This model proposes that the awareness that one's fellow in-group members have close friendships with out-group members can help to reduce prejudice toward the out-group. A likely reason for this is that the in-group "exemplars" provide normative information regarding how one should behave and also may contribute to the redefinition of the intergroup relationship as less negative. Such an "extended contact effect" does not require personal intergroup friendship for the perceiver. Correlational and experimental evidence of this idea is emerging. Underlying this model is the idea that both in-group and out-group role models should be seen as typical or representative of their groups and not as exceptions to the rule.

7.3. The Decategorization and Recategorization Models for Reducing Intergroup Bias

The so-called decategorization model of intergroup relations holds that intergroup contact is more effective when group salience is low and the intergroup situation is personalized. However, individual-to-group generalizations are unlikely in very personalized contact situations. In contrast, the so-called common in-group identity model holds that intergroup attitudes are best improved when the category boundaries are drawn so that the out-group becomes subsumed into a new and larger superordinate category. There is now an impressive literature supporting this claim. However, most studies of this model have not examined generalization. The fundamental limitation of both the decategorization and recategorization models is that they threaten to deprive individuals of valued social identities in smaller, less inclusive groups. A common in-group identity may be only short-lived or unrealistic in the face of powerful ethnic/racial categorizations (e.g., the former Yugoslavia). In addition, for groups with histories of antagonism and for minorities whose members are likely to resist assimilation into a superordinate category that is dominated by a majority out-group, the prospect of a superordinate identity may constitute a threat that actually increases bias.

7.4. Integrating Different Models of Contact

To overcome these problems, various solutions have been proposed. Hewstone and colleagues in 2002 proposed that the categorization salience model of contact needs to be integrated with the more personalized models and that, in contrast to original social identity theory, interpersonal and intergroup contact should be viewed as orthogonal dimensions that together can create highly effective conditions of out-group contact (i.e., contact should simultaneously be both highly intergroup and highly interpersonal).

The dual-identity model developed by Gaertner and colleagues takes this idea as a starting point as it aims to maximize the benefits of both the "common in-group identity" and the mutual intergroup differentiation model. According to this model, subgroup identity is maintained while the groups also share membership in a common superordinate group. A dual identity leads to more positive out-group attitudes than does a superordinate identity alone. Gaertner and colleagues conducted a correlational study in a multiethnic high school and found less bias when students identified themselves at both the subgroup (i.e., ethnic group) and superordinate (i.e., American) levels rather than at the subgroup level only.

The dual-identity model resembles the bidimensional interactive acculturation model suggested by Bourhis and colleagues. According to this model, as well as to the original acculturation theory developed by Berry during the 1980s, members of cultural minorities can adopt different acculturation strategies. These strategies are formed by the crossing of two orthogonal orientations: in the interactive acculturation model, these are defined as identification with one's own culture (high/low) and identification with the dominant culture (high/low). Research suggests that the "intergration" strategy (high on both orientations) is often the preferred strategy among minority groups and seems to be associated with less acculturative stress and better educational outcomes.

A major problem for the dual-identity approach is that members of minority and majority groups may have different preferences for what model of intergroup relations to adopt. Dominant majority ethnic groups tend to favor assimilation, whereas racial/ethnic minorities favor pluralistic integration. Thus, dual identity may reduce bias for the minority group but not for the majority group. This asymmetry between minority and majority strategies is also reflected in the interactive acculturation model itself in that the strategies of majority members concern only their preferences for the identification strategies they would like the minority members to adopt. As a consequence, minority and majority members can, at best, agree only on the most preferred strategy for the minority group members. Even then, majority members might not be prepared to change their own identification patterns accordingly.

A successful superordinate category and identity must be inclusive rather than reflect the superior characteristics of a dominant majority group too strongly. Even if this is the case, however, subgroup categorizations and identities are likely to be stronger and more stable than superordinate ones. One simple way in which to reduce intergroup bias is to ensure that the contact partners can be simultaneously classified as in-group and out-group members on multiple dimensions. Shared or overlapping category memberships reduce bias simply because they make social categorization more complex and decrease the importance of any one in-group/out-group distinction. However, the effectiveness of crossed categorization is limited when one category distinction is functionally dominant and/or categories are correlated, and this is normally the case in real-life intergroup encounters.

All of the models for reduction of intergroup bias presented here should be seen as complementary and reciprocal rather than as competing and exclusive. This integrative approach responds to the fact that each model can be effective under particular contact conditions but that each model also has its weaknesses and limitations, particularly with respect to generalization and to finding an intervention that works for both majority and minority groups.

One way in which to combine the various models was proposed by Pettigrew in 1998. He suggested that time sequence is crucial. Diminished category salience can be important when intergroup contact is initiated, but category salience may be required later for the positive effects to generalize to the intergroup level. At an even later stage, the recategorization strategy proposed by the common in-group identity model may become possible. After continuous contact, people may begin to think of in-group and out-group members as members in a more inclusive common in-group, although this final state is seldom reached by interacting groups.

7.5. Realistic Goals for Real-Life Intergroup Relations

Prior attitudes and experiences influence whether people seek or avoid intergroup contact and what the effects of the contact will be. High intergroup anxiety and threat can impede both contact and its positive effect. In addition, institutional and societal norms structure the form and effects of contact situations. Optimal contact situations are difficult to achieve if a struggle for power fuels the larger intergroup conflict. In contrast, normative support makes attainment of optimal conditions much easier. Although the beneficial effects of contact with typical out-group members have been confirmed in a number of studies, if the intergroup encounter goes wrong, structuring the encounter as intergroup rather than interpersonal could make matters worse. Therefore, there is a fine line to be drawn in designing situations of intergroup contact. Enough category salience must be retained to permit the positive change to generalize, but not so much as to allow the regression into familiar patterns of prejudice that are associated with category divisions that are too firmly drawn.

Given the corrosive nature of many real-world conflicts, out-group liking is often unlikely; and achieving other outcomes may be more realistic and just as important, including increases in perceived group variability, out-group knowledge, and perspective taking as well as decreases in intergroup anxiety. To reduce full-blown intergroup conflict, effective interventions also need to build trust. It is true that relevant research and theoretical developments on intergroup bias have been directed at its relatively mild forms. However, social categorizations play a role in even extreme forms of prejudice. Severely prejudiced individuals and ideologies delegitimize victims, assigning them to an extreme social category that enjoys no projection, and dehumanize victims by morally excluding them from the in-group of humanity where justice, fairness, and morality prevail.

However, real-world intergroup relations owe at least as much to intergroup history, economics, politics, and ideology as they do to social psychological variables such as categorization, self-esteem, in-group identifications, group size, and group threat. Social conflict is more complex than intergroup bias. Consequently, intergroup bias should not be analyzed from just one disciplinary

perspective, regardless of how important the psychological antecedents and consequences of intergroup bias may be. The current trend to study culture in social psychology has yet to address intergroup bias in full force, but there is little doubt that it will. Culture and societal influences are neglected at our peril.

See Also the Following Articles

Motives and Goals ■ Prejudice and Discrimination ■ Stereotypes

Further Reading

Brown, R. (2001). Intergroup relations. In M. Hewstone, & W. Stroebe (Eds.), *Introduction to social psychology* (pp. 480–515). Oxford, UK: Blackwell.

Fiske, S. (2000). Stereotyping, prejudice, and discrimination as seen between the centuries: Evolution, culture, mind, and brain. *European Journal of Social Psychology*, *30*, 299–322.

Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual Review of Psychology*, *53*, 575–604.

Liebkind, K. (2001). Acculturation. In R. Brown, & S. Gaertner (Eds.), *Blackwell handbook of social psychology: Intergroup processes* (pp. 386–406). Oxford, UK: Blackwell.

Pettigrew, T. F. (1998). Intergroup contact theory. *Annual Review of Psychology*, *49*, 65–85.

Smith, P. B., & Bond, M. H. (1999). *Social psychology across cultures* (2nd ed.). Boston: Allyn & Bacon.

Triandis, H. C., & Trafimov, D. (2001). Culture and its implication for intergroup behavior. In R. Brown, & S. Gaertner (Eds.), *Blackwell handbook of social psychology: Intergroup processes* (pp. 367–384). Oxford, UK: Blackwell.

Verkuyten, M. (2000). The benefits to social psychology of studying ethnic minorities. *European Bulletin of Social Psychology*, *12*(3), 5–21.

International Classification of Diseases (WHO)

Jesús Rodríguez-Marín
Miguel Hernandez University, Alicante, Spain

1. Introduction
2. International Statistical Classification of Diseases and Related Health Problems, 10th Revision
3. ICD-10's Chapter 5: Mental and Behavioral Disorders
4. Phenomenological Rigor, Consistency, and Diagnostic Stability of ICD-10 Psychiatric Diagnoses
5. Clinical Utility of ICD-10
6. Comparison between ICD-10 and DSM-IV
7. Collaborating Centers for Classification of Diseases and Procedures for ICD-10 Updating
Further Reading

The *International Statistical Classification of Diseases* (ICD) was worked out and proposed by the World Health Organization (WHO) to facilitate communication and study on organic and mental diseases among health workers and clinicians worldwide. In 1893, Jacques Bertillon, head of statistical works for the city of Paris, presented a new classification of *Diseases and Causes of Death* to the International Congress of Statistics in Chicago. The Bertillon classification was the first of a series that gave rise to the ICD, published by WHO from 1948 onward, to improve diagnosis and classification of diseases and to promote international comparability in the collection, processing, classification, and presentation of mortality and morbidity statistical data. Since then, the ICD has been revised periodically to incorporate changes in the medical field. This article includes a historical approach to the ICD, and its general principles and structure. The article specifically discusses Chapter 5 of the 10th revision of the ICD (ICD-10), on mental and behavioral disorders, describing the phenomenological rigor, consistency, diagnostic stability, and clinical utility of the psychiatric diagnoses. Finally, it incorporates a brief comparison of ICD-10 and the 4th edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV) and mentions WHO collaborating centers for classification of diseases and procedures for ICD-10 updating.

1. INTRODUCTION

International collaboration began a century ago with the aim of drawing up a classification of diseases. The systematic classification of diseases began in the 18th century with the publication of the *Nosologia Methodica* by Lacroix and of the *Genera Morborum* by Linnaeus. In 1785, Cullen published the *Synopsis Nosologiae Methodicae,* which was used in various European countries until the middle of the 19th

©2004 Elsevier Inc.
All rights reserved.

century. A classification of diseases and death causes was approved in 1855 during the Second International Congress of Statistics in Paris. This classification, which combined the English tradition (based on anatomical region) and the French tradition (based on etiology), had wide repercussions but did not garner complete acceptance.

In 1893, Jacques Bertillon, head of statistical works for the city of Paris, presented a new classification of *Diseases and Causes of Death,* based on the one employed in Paris, to the International Congress of Statistics in Chicago. The Congress approved it and recommended its use and revision every 10 years. The Bertillon classification of diseases was the first of a series that gave rise to the *International Statistical Classification of Diseases* (ICD), edited and published by the World Health Organization (WHO) from 1948 onward, to improve diagnosis and classification of diseases and to promote international comparability in the collection, processing, classification, and presentation of mortality and morbidity statistical data. Since then, the ICD has been revised periodically to incorporate changes in the medical field. Bertillon himself did the work in 1900, 1910, and 1920. WHO assumed the task from the 4th revision (in 1929) onward. The 4th revision and the next one (in 1938) maintained the basic structure of the Bertillon classification. In both of these, the mental diseases were included without any special section; instead, they were distributed throughout 9 of the 14 sections.

In 1948, WHO's 6th revision included a chapter specifically dedicated to mental diseases (Chapter 5), and this remained unchanged until the 9th revision (ICD-9). This new chapter contained three sections dedicated to (a) psychosis, (b) psychoneurotic disorders, and (c) disorders of character, behavior, and intelligence. The 1955 edition dedicated more space to mental disorders but maintained these three sections.

Since the beginning of the 1960s, WHO has stimulated a series of works on mental disorders, supported projects to improve diagnosis and classification of mental diseases, and encouraged research on classification instruments. Emphasis was placed on the reliability, consistency, and diagnostic stability of classification instruments. These efforts culminated in a program with four main subprograms dedicated to (a) the improvement of the diagnosis and classification of psychiatric disorders, (b) cross-cultural research of the most important psychiatric disorders, (c) an epidemiological study, and (d) training in social and psychiatric epidemiology. A network of individuals and institutions was created, and they worked on proposals to be included in the draft for the 8th ICD revision (in 1968). In this 8th revision, the diagnostic categories for mental diseases were reorganized into three sections: (a) psychosis; (b) neurosis, personality disorders, and other nonpsychotic mental disorders; and (c) mental retardation. A glossary defined each diagnostic category and standardized its use.

The 9th revision of the ICD (in 1978) contained four sections dedicated to mental disorders: (a) organic psychosis; (b) other psychoses; (c) neurotic disorders, personality disorders, and other nonpsychotic mental disorders; and (d) mental retardation. The difficulties of diagnosis and classification of some nosologic entities revived interest and research into psychiatric diagnosis, and WHO again took on the leadership in the preparation of the 10th ICD edition.

The rules and study areas for classification of mental disturbance were defined during an important conference in Copenhagen in 1982. Also identified were those areas in which improvement in classification required further research work and more evidence. A trial classification was then staged in 55 countries with the participation of more than 700 clinicians from at least 100 centers, coordinated by a group of 13 centers. More than 15,000 patients were assessed in the trial classification.

Some special notes should be made here with regard to the development of disease classification in the 10th revision (ICD-10) in 1992. First, it was developed simultaneously in several languages, and it enabled compatible definitions among many countries. Second, the different versions of ICD-10 allowed different uses in accordance with different necessities. Third, the classification was based on the evidence found in research. Fourth, the classification was accompanied by standard instruments for the evaluation of symptoms and other diagnosis elements.

The results of this project were very positive.

2. INTERNATIONAL STATISTICAL CLASSIFICATION OF DISEASES AND RELATED HEALTH PROBLEMS, 10th REVISION

In the 10th revision, the title was changed to clarify the content and goals and to reflect the progressive extension of the scope of the classification.

ICD-10 is much longer than the 9th revision. Printed in three volumes (the 9th revision was printed in two volumes), ICD-10 has adopted an alphanumeric coding scheme (rather than using numeric categories) and has significantly enlarged the number of categories available for classification. The first character in the coding system is now a letter for each chapter, as well as a number, and each chapter is assigned 10 two-character categories and 100 three-character major categories. ICD-10 has nearly twice as many categories as did the 9th revision. Finally, some slight changes have been made in the coding rules for mortality. For instance, in the case of Chapter 5 (on mental and behavioral disorders), the three-character categories go from F00 to F99. F20 is "schizophrenia," and every category has different subcategories (e.g., F20.0 is "paranoid schizophrenia," F20.5 is "residual schizophrenia").

The first volume of ICD-10 contains 17 chapters dedicated to different diseases (from "certain infectious and parasitic diseases" to "congenital malformations, deformations, and chromosomal abnormalities"), special tabulation lists for mortality and morbidity, the definitions, and the nomenclature regulations. The second volume includes new background information, instructions, and guidelines for users of the tabular list and historical information about the development of the ICD classification. The third volume incorporates a detailed alphabetical index. The index includes diagnostic terms commonly used as synonyms for the terms officially accepted in the classification, and it facilitates more efficient coding.

The ICD-10 manual includes not only the disorders but also other attributes of patients' interactions with their social contexts. Consequently, it includes a multiaxial system with three axes, with Axis I covering clinical diagnoses of both mental and physical disorders, Axis II covering disabilities due to impairments produced by the disorders, and Axis III covering the contextual or environmental and personal lifestyle factors influencing the appearance and course of the disorders. The ICD-10 manual is accompanied by the ICD-10 Multiaxial Diagnostic Formulation Form, the WHO Short Disability Assessment Schedule (WHO DAS-S), and Axis I and Axis III glossaries.

ICD-10 is intended for use in clinical, educational, and research activities. It has been prepared by an international team of experts and has been field-tested in 20 countries. It has been shown to be easy to use and applicable in a wide range of cultures and settings.

3. ICD-10'S CHAPTER 5: MENTAL AND BEHAVIORAL DISORDERS

ICD-10 dedicates Chapter 5 to mental and behavioral disorders. This chapter, under the letter F of the coding system, substantially extends the space given previously to mental disorders.

Chapter 5 in ICD-10 offers important differences compared with the previous revision. That chapter in the 9th revision had only 30 three-character categories (290 to 319), whereas it includes 100 major three-character categories in ICD-10. Two other important changes have been made. First, differentiation between psychoses and neuroses has been abandoned. Second, disorders showing common elements or properties are grouped together.

Chapter 5 in ICD-10 contains the following sections: F00 to F09, organic (including symptomatic) disorders; F10 to F19, mental and behavioral disorders due to psychoactive substances; F20 to F29, schizophrenia, schizotypal, and delusional disorders; F30 to F39, mood (affective) disorders; F40 to F49, neurotic, stress-related, and somatoform disorders; F50 to F59, behavioral syndromes associated with physiological disturbances and physical factors; F60 to F69, disorders of adult personality and behavior; F70 to F79, mental retardation; F80 to F89, disorders of psychological development; F90 to F98, behavioral and emotional disorders with onset usually occurring during childhood or adolescence; and F99, unspecified mental disorders.

The most recent Chapter 5 also contains clinical guidelines for the classification of mental disorders in several languages; compatible versions dedicated to classification, clinical, and investigation work; and a multiaxial presentation with three axes: clinical diagnosis, the development of factors related to disorders, and the disability level related to them. In addition, ICD-10 includes a version for primary care doctors.

The publication of the ICD-10 classification of mental and behavioral disorders has been an extremely important event for both psychiatry and clinical psychology. For the first time, psychiatry has produced a classification based on the worldwide consensus of experts and schools of psychiatry. ICD-10 constitutes an essential basis that enables the development of a unique language for national and international communication in the field of mental health, facilitating collaboration across cultures and countries.

Chapter 5 of ICD-10 has been translated into more than 35 languages and presented in different versions for

different purposes. Everywhere, the code numbers and titles of the disorders are exactly the same, but details and styles of presentation are adapted to specific purposes. Furthermore, two major documents have been developed from ICD-10: a version for general clinical practice and educational use (*ICD-10 Clinical Descriptions and Diagnostic Guidelines*) and another version for research purposes (*ICD-10 Diagnostic Criteria for Research*).

4. PHENOMENOLOGICAL RIGOR, CONSISTENCY, AND DIAGNOSTIC STABILITY OF ICD-10 PSYCHIATRIC DIAGNOSES

ICD-10 Clinical Descriptions and Diagnostic Guidelines and *ICD-10 Diagnostic Criteria for Research* have been analyzed for their internal consistency and phenomenological soundness. The operational definitions of all categories have been considered both for their agreement with the core text of ICD-10 and for their internal consistency within each category. The soundness of all operational criteria and concepts used in the clinical descriptions has also been assessed. The results indicate a high degree of consistency and phenomenological rigor as well as a high level of interrater reliability. Nonetheless, some inconsistencies have been detected in a few categories related to mania, hypomania, cyclothymia, schizotypal disorder, and behavior disorder. Especially in the schizotypal disorder category, a low interrater reliability and poor face validity were found, and this could lead to the label of mental illness being applied to otherwise healthy people.

Temporal stability of diagnosis, a measure of the degree to which diagnoses remain unchanged at later hospital admissions, seems to be high for schizophrenic and affective disorders, whereas it is not so high for neurotic, obsessive–compulsive, stress-related, and adjustment disorders. Further research is required in this area.

5. CLINICAL UTILITY OF ICD-10

ICD-10 has led to a major innovation in clinical psychiatry. The criteria-based classifications are no longer simple coding conventions; instead, they have become part of the conceptual framework of the discipline itself. Consequently, comprehensive clinical evaluation should work at two separate levels: (a) the syndrome diagnosis and (b) the evaluation of factors possibly influencing the course and outcome of psychiatric disorders. Only when both are taken into consideration is it possible to get a complete understanding of the disorders and optimal treatment.

6. COMPARISON BETWEEN ICD-10 AND DSM-IV

During the 1940s, the American Psychiatric Association drew up a psychiatric classification system of mental diseases titled *Diagnostic and Statistical Manual of Mental Disorders* (DSM), and since then it has been the main reference for contrasting the ICD. The DSM is produced at intervals of at least 10 years, reflects U.S. national trends and opinions, and has probably been the most important of the national classifications. Therefore, although ICD-10 is the official coding system in many countries, DSM-IV appears to be more popular among mental health professionals.

Both systems, ICD-10 in the *Diagnostic Criteria for Research* and the 4th edition of the DSM (DSM-IV), have introduced descriptive, nonetiological classifications with operationally defined categories based on diagnostic criteria and have proposed explicit diagnostic criteria. There are differences between the two systems, with the most significant ones being the avoidance of social dysfunction criteria in ICD-10 and the hierarchical order preventing comorbidity diagnoses in ICD-10. Results from separate studies show high levels of concordance in diagnoses of depression, dysthymia, substance dependence, and generalized anxiety disorder. They also show a moderate concordance in social phobia, obsessive–compulsive disorder, and the three panic/agoraphobia disorders. Results show only a low concordance for posttraumatic stress disorder and drug use/abuse. Concordances and differences basically reflect similarities and dissimilarities in concepts and diagnostic criteria between classification systems.

Both systems have multiaxial presentations, but it is doubtful that the ICD-10 multiaxial presentation could be used to the same extent as can the DSM-IV multiaxial assessment. ICD-10 has sought to place psychiatric disorders in the context of the different religions, nationalities, and cultures of the world community, whereas DSM-IV lacks this degree of cultural sensitivity, perhaps because it is a system that originates from only one country. The concern for cultural universality increased the length of the document but curtailed and

simplified some operational definitions by omitting culturally biased criteria. However, ICD-10 does contain some questions that are not culturally sensitive because it is based on premises that are inherent in its construction. Consequently, and in spite of efforts to coordinate ICD-10 and DSM-IV as much as possible, major differences in important areas remain, although the two systems use the same coding system in some cases.

7. COLLABORATING CENTERS FOR CLASSIFICATION OF DISEASES AND PROCEDURES FOR ICD-10 UPDATING

To assist users with problems in the development and use of disease classifications and in the use of the ICD, 10 WHO Collaborating Centers for the Classification of Diseases have been created: 3 centers for English-language users (in London, United Kingdom; Canberra, Australia; and Hyattsville, Maryland, United States), 1 center for French- users (Le Vésinet, France), 1 center for Arabic users (Safat, Kuwait), 1 center for Chinese users (Beijing, China), 1 center for Nordic users (Uppsala, Sweden), 1 center for Portuguese users (Sao Paulo, Brazil), 1 center for Russian users (Moscow, Russia), and 1 center for Spanish users (Caracas, Venezuela).

Proposals for changes or additions to ICD-10 should first be presented to the relevant collaborating center. All proposals for changes must be sponsored by, and submitted to WHO through, a collaborating center. Different meetings of WHO Collaborating Centers for Classification of Diseases have approved specific procedures for updating ICD-10.

See Also the Following Articles

Cognitive and Behavioral Interventions for Persons with Dementia ■ Dementia in Older Adults ■ Diagnostic and Statistical Manual of Mental Disorders

Further Reading

Andrews, G., & Slade, T. (1998). Depression, dysthymia, and substance use disorders: Sources of dissonance between ICD-10 and DSM-IV. *International Journal of Methods in Psychiatric Research*, *7*, 116–120.

Andrews, G., Slade, T., & Peters, L. (1999). Classification in psychiatry: ICD-10 versus DSM-IV. *British Journal of Psychiatry*, *174*, 3–5.

Bertelsen, A. (1999). Reflections on the clinical utility of the ICD-10 and DSM-IV classifications and their diagnostic criteria. *Australian and New Zealand Journal of Psychiatry*, *33*, 166–173.

Cooper, J. E. (1988). The structure and presentation of contemporary psychiatric classifications with special reference to ICD-9 and -10. *British Journal of Psychiatry*, *152*(Suppl.), 21–28.

Haghighat, R. (1997). Phenomenological rigour and consistency of ICD-10. *European Journal of Psychiatry*, *11*, 90–93.

Haghighat, R. (1999). Cultural sensitivity: ICD-10 versus DSM-III-R. *International Journal of Social Psychiatry*, *40*, 189–193.

Rice, J. P., Rochberg, N., Endicott, J., Lavori, P. W., & Miller, C. (1992). Stability of psychiatric diagnoses: An application to the affective disorders. *Archives of General Psychiatry*, *49*, 824–830.

Sartorius, N. (1991). The classification of mental disorders in the tenth revision of the *International Classification of Diseases*. *European Psychiatry*, *6*, 315–322.

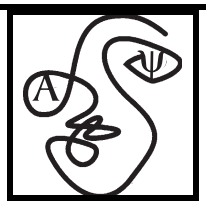
World Health Organization. (1992a). *ICD-10 clinical descriptions and diagnostic guidelines*. Geneva, Switzerland: Author.

World Health Organization. (1992b). *International statistical classification of diseases and related health problems* (10th rev.). Geneva, Switzerland: Author.

World Health Organization. (1993). *ICD-10 diagnostic criteria for research*. Geneva, Switzerland: World Health Organization.

World Health Organization. (1996). *WHO ICD-10 diagnostic and management guidelines for mental disorders in primary care (ICD-10, Chapter 5, primary care version)*. Seattle, WA: Hogrefe & Huber.

World Health Organization. (1997). *The multiaxial presentation of the ICD-10 for use in adult psychiatry*. Cambridge, UK: Cambridge University Press.



International Conflict

Herbert C. Kelman
Harvard University, Cambridge, Massachusetts, USA

1. International Conflict as a Contemporary Problem
2. Social–Psychological Propositions about the Nature of International Conflict
3. Psychological Processes Promoting Conflict
4. Implications for Conflict Resolution
Further Reading

GLOSSARY

attribution mechanisms Ways in which to assess the motivation for the behavior of self and others; in conflict relationships, these often lead to confirmation of the negative attitudes toward the adversary.

coalition across conflict lines A working relationship between members of opposing parties seeking common ground in the pursuit of their respective interests.

conflict A relationship between two parties characterized by the fact and/or perception that each party, in the pursuit of its own interests, threatens or undermines the interests of the other party.

conflict norms A set of social expectations within groups or societies in conflict that require members to manifest hostility and distrust toward the adversary in their feelings, language, and actions.

enemy image The tendency among parties in conflict to demonize and delegitimize the adversary and to perceive the other party as inherently aggressive, dangerous, and inauthentic.

groupthink A tendency within groups tasked with decision making in crisis situations to seek concurrence among their members, discourage dissenting views, and thereby maintain group cohesiveness.

human needs approach A view of conflict as generated and maintained by the nonfulfillment and threat to the parties' basic human needs and of conflict resolution as a process that addresses the basic needs of both parties.

integrative solutions "Win–win" solutions to a conflict that meet the needs and interests of both parties; these are in contrast to "win–lose" solutions, in which one party prevails at the expense of the other.

interactive problem solving An unofficial approach to conflict resolution in which the parties are encouraged to treat the conflict as a shared problem that requires joint efforts to shape mutually satisfactory solutions.

mirror image formation The tendency among parties in conflict to develop parallel images of self and the enemy, except with the values reversed, for example, for each party to see the other as inherently aggressive and the self as entirely defensive in response.

mutual reassurance A form of influence in conflict relationships designed to demonstrate each party's awareness of the other's concerns and requirements and thereby to reduce the other party's fears of the risks entailed by negotiation.

negotiation of identity An informal process whereby parties in conflict explore and invent ways in which to accommodate their group identities and national narratives to one another, particularly by removing negation of the other from their own identities and narratives.

prospect theory A theory of decision making under conditions of risk that proposes that people are more prone to take risks to avoid losses than to achieve gains, for example, to go to war in defense of threatened interests than to pursue a potential opportunity for peace.

realistic empathy Taking the perspective of the adversary to gain a realistic understanding of the other party's thoughts and feelings.

©2004 Elsevier Inc.
All rights reserved.

responsiveness An influence strategy based on attempts to elicit cooperation from the other party by addressing the other's needs, fears, and constraints.

self-fulfilling prophecy A relationship tendency whereby one party's expectations of the other cause the other party to behave in ways that confirm these expectations.

zero-sum thinking A view of the conflict as a relationship where every gain by the other party represents a loss for one's own side and where every loss by the other party represents a gain for one's own side.

The end of the cold war has changed the character of international conflict, but it has not reduced the prevalence of violent conflict, the human suffering that it causes, or the threat that it poses to the well-being and survival of the human species. Thus, during recent years, there has been a proliferation of deadly, deep-rooted conflicts between ethnic and other identity groups within and across nation-states—conflicts often marked by violence against civilians, ethnic cleansing, and genocidal actions. There has also been a rise in terrorism and counterterrorism around the world, undermining the peace and development of established and emerging states.

1. INTERNATIONAL CONFLICT AS A CONTEMPORARY PROBLEM

Psychological theory and research can make useful contributions to the understanding and amelioration of the violent manifestations of international conflict. Indeed, psychological concepts and findings have been used increasingly in the study of international conflict and international relations more generally as well as in the development of new approaches to conflict resolution. This article offers a perspective on international conflict that is anchored in social–psychological theory and research and that, in turn, informs the practice of conflict resolution. Social–psychological analysis is designed to complement (and not to replace) approaches based on structural or strategic analysis by providing a special lens for viewing international conflict that brings some of its less explored dimensions into focus.

In this spirit, the article begins with a discussion of several propositions about the nature of international conflict that flow from a social–psychological perspective and that have clear implications for conflict resolution. It then describes social–psychological processes characteristic of conflict interaction that contribute to the escalation and perpetuation of conflict and that must be reversed if such conflict is to be resolved. [These ideas are presented in greater detail in Kelman's chapter in *Peacemaking in International Conflict* (edited by Zartman and Rasmussen), and linked to a discussion of approaches to conflict resolution in Kelman and Fisher's chapter in the *Oxford Handbook of Political Psychology* (edited by Sears, Huddy, and Jervis).]

2. SOCIAL–PSYCHOLOGICAL PROPOSITIONS ABOUT THE NATURE OF INTERNATIONAL CONFLICT

A social–psychological perspective can expand on and enrich the analysis of international conflict provided by the realist or neo-realist schools of international relations or other, more traditional approaches. While acknowledging the importance of objectively anchored national interests, it explores the subjective factors that set constraints on rationality. Without denying the primacy of the state in the international system, a social–psychological perspective opens the "black box" of the state as unitary actor and analyzes processes within and between the societies that underlie state action. Fully recognizing the role of power in international relations, it postulates a broader range of influence processes (and, indeed, of definitions of power) that play a role in international politics. And while affirming the importance of structural factors in determining the course of an international conflict, it conceives of conflict as a dynamic process, shaped by changing realities, interests, and relationships between the conflicting parties.

These observations suggest four general propositions about international conflict that call for social–psychological concepts and data. The four propositions are particularly relevant to existential conflicts between identity groups, that is, conflicts in which the collective identities of the parties are engaged and in which the continued existence of the groups is seen to be at stake. Thus, the propositions apply most directly to ethnic or ideological conflicts, but they also apply to more mundane interstate conflicts insofar as issues of national identity and existence come into play—as they often do.

2.1. Conflict as a Process Driven by Collective Needs and Fears

According to the first proposition, international conflict is a process driven by collective needs and fears

rather than entirely a product of rational calculation of objective national interests on the part of political decision makers. Human needs are often articulated and fulfilled through important collectivities such as the ethnic group, the national group, and the state. Conflict arises when a group is faced with nonfulfillment or threat to the fulfillment of basic needs, not only obvious material needs such as food, shelter, physical safety, and physical well-being but also psychological needs such as identity, security, recognition, autonomy, self-esteem, and a sense of justice—needs featured in the 1990 volume *Conflict* and other writings by Burton, a pioneer in the field of conflict resolution. Moreover, needs for identity and security and similarly powerful collective needs, as well as the fears and concerns about survival associated with them, contribute heavily to the escalation and perpetuation of conflict once it has started. Even when the conflicting parties have come to the conclusion that it is in their best interest to put an end to the conflict, they resist going to the negotiating table or making the accommodations necessary for the negotiations to move forward out of fear that they will be propelled into concessions that, in the end, will leave their very existence compromised. The fears that drive existential conflicts lie at the heart of the relationship between the conflicting parties, going beyond the cycle of fears resulting from the dynamics of the security dilemma discussed by Jervis in 1976.

Collective fears and needs, although more pronounced in ethnic conflicts, play a part in all international conflicts. They combine with objective factors—a state's resources, the ethnic composition of its population, its access to the sea (or lack thereof), and the like—in determining how different segments of a society perceive state interests and what ultimately becomes the national interest as defined by the dominant elites. Similarly, all conflicts—interstate as well as ethnic—represent a combination of rational and irrational factors, and in each type of conflict the mix may vary from case to case. Some ethnic conflicts may be preponderantly rational, just as some interstate conflicts may be preponderantly irrational. Furthermore, in all international conflicts, the needs and fears of populations are mobilized and often manipulated by the leadership, with varying degrees of demagoguery and cynicism. Even when manipulated, collective needs and fears represent authentic reactions within the population and become the focus of societal action. They may be linked to individual needs and fears. For example, in highly violent ethnic conflicts, the fear of annihilation of one's group is often tied to a fear of personal annihilation (and for good reason).

The conception of conflict as a process driven by collective needs and fears implies, first and foremost, that conflict resolution—if it is to lead to a stable peace that both sides consider just and to a new relationship that enhances the welfare and development of the two societies—must address the fundamental needs and deepest fears of the populations. From a normative point of view, such a solution can be viewed as the operationalization of justice within a problem-solving approach to conflict resolution. Another implication of a human needs orientation noted by Burton is that the psychological needs on which it focuses—security, identity, recognition, and the like—are not inherently zero-sum, although they are usually seen as such in deep-rooted conflicts. Thus, it may well be possible to shape an integrative solution that satisfies both sets of needs, which in turn might make it easier to settle issues such as territory and resources through distributive bargaining. Finally, the view of conflict as a process driven by collective needs and fears suggests that conflict resolution must, at some stage, provide for certain processes that take place at the level of individuals and interactions between individuals, for example, taking the other society's perspective (what White called "realistic empathy" in 1984 and elsewhere) as well as creative problem solving, learning, and insight.

2.2. Conflict as an Intersocietal Process

Focusing on the needs and fears of the populations in conflict reminds one that international conflict is an intersocietal process, not merely an intergovernmental or interstate phenomenon. The conflict, particularly in the case of protracted ethnic struggles, becomes an inescapable part of daily life for each society and its component elements. Thus, analysis of conflict requires attention, not only to its strategic, military, and diplomatic dimensions but also to its economic, psychological, cultural, and social–structural dimensions. Interactions along these dimensions, both within and between the conflicting societies, shape the political environment in which governments function and define the political constraints under which they operate.

An intersocietal view of conflict points to the role of internal divisions within each society, which often play a major part in exacerbating or even creating conflicts between the societies. They impose constraints on

political leaders pursuing a policy of accommodation in the form of accusations by opposition elements that the leaders are jeopardizing national existence as well as in the form of anxieties and doubts within the general population that are both fostered and exploited by the opposition elements. However, the internal divisions may also provide potential levers for change in the direction of conflict resolution by challenging the monolithic image of the enemy that parties in conflict tend to hold and by enabling them to deal with each other in a more differentiated way. They point to the presence of potential partners for negotiation on the other side and, thus, provide the opportunity for forming what Kelman called pro-negotiation "coalitions across the conflict lines" in 1993. To contribute to conflict resolution, any such coalition must necessarily remain an "uneasy coalition," lest its members lose their credibility and political effectiveness within their respective communities.

Another implication of an intersocietal view of conflict is that negotiations and third-party efforts should ideally be directed not to mere settlement of the conflict, in the form of a brokered political agreement, but to its resolution. A political agreement may be adequate for terminating relatively specific, containable interstate disputes, but conflicts that engage the collective identities and existential concerns of the adversaries require a process conducive to structural and attitude change, to reconciliation, and to the transformation of the relationship between the two societies. Finally, an intersocietal analysis of conflict suggests a view of diplomacy as a complex mix of official and unofficial efforts with complementary contributions. The peaceful termination or management of conflict requires binding agreements that can be achieved only at the official level, but many different sectors of the two societies must be involved in creating a favorable environment for negotiating and implementing such agreements.

2.3. Conflict as a Multifaceted Process of Mutual Influence

International conflict is best understood as a multifaceted process of mutual influence and not just a contest in the exercise of coercive power. Much of international politics entails a process of mutual influence in which each party seeks to protect and promote its own interests by shaping the behavior of the other party. Conflict occurs when these interests clash, that is, when attainment of one party's interests (and fulfillment of the needs that underlie them) threatens, or is perceived to threaten, the interests (and needs) of the other party. Therefore, in pursuing the conflict, the parties engage in mutual influence, designed to advance their own positions and to block the adversary. Similarly, in conflict resolution, by negotiation or other means, the parties exercise influence to induce the adversary to come to the table, to make concessions, to accept an agreement that meets their interests and needs, and to live up to that agreement. Third parties also exercise influence in conflict situations by backing one party or the other, by mediating between them, or by maneuvering to protect their own interests.

Influence in international conflict typically relies on a mixture of threats and inducements, with the balance often on the side of force and the threat of force. Thus, the U.S.–Soviet relationship during the cold war was predominantly framed in terms of an elaborate theory of deterrence, that is, a form of influence designed to keep the other side from doing what one does not want it to do, as described in important books by Schelling in 1963, by George and Smoke in 1974, and by Jervis and colleagues in 1985. In other conflict relationships, the emphasis may be on compellence, that is, a form of influence designed to make the other side do what one wants it to do. Such coercive strategies entail serious costs and risks, and their effects may be severely limited. For example, they are likely to be reciprocated by the other side and lead to escalation of the conflict, and they are unlikely to change behavior to which the other side is committed. Thus, the effective exercise of influence in international conflict requires a broadening of the repertoire of influence strategies, at least to the extent of combining "carrots and sticks," that is, of supplementing the negative incentives that typically dominate international conflict relationships with positive incentives (e.g., economic benefits, international approval, a general reduction in the level of tension). An example of an approach based on the systematic use of positive incentives is Osgood's GRIT (Graduated and Reciprocated Initiatives in Tension reduction) strategy. Egyptian President Anwar Sadat, in his 1977 trip to Jerusalem, undertook a unilateral initiative with the expectation (partly prenegotiated) of Israeli reciprocation, but—unlike GRIT—he started with a large fundamental concession in the anticipation that negotiations would fill in the intervening steps.

Effective use of positive incentives requires more than offering the other party whatever rewards, promises, or confidence-building measures seem to be most readily available. It requires actions that address

the fundamental needs and fears of the other party. Thus, the key to an effective influence strategy based on the exchange of positive incentives is responsiveness to the other party's concerns, that is, actively exploring ways in which each party can help to meet the other's needs and allay its fears as well as ways in which the parties can help each other to overcome the constraints within their respective societies against taking the actions that each wants the other to take. The advantage of a strategy of responsiveness is that it allows each party to exert influence on the other through positive steps (not threats) that are within its own capacity to take. The process is greatly facilitated by communication between the parties to identify actions that are politically feasible for each party and yet are likely to have an impact on the other.

A key element in an influence strategy based on responsiveness is mutual reassurance, which is particularly critical in any effort to resolve an existential conflict. The negotiation literature suggests that parties are often driven to the table by what Zartman in 1989 described as "a mutually hurting stalemate," which makes negotiations more attractive than continuing the conflict. But parties in existential conflicts are afraid of negotiations, even when the status quo has become increasingly painful and they recognize that a negotiated agreement is in their interest. To advance the negotiating process under such circumstances, it is at least as important to reduce the parties' fears as it is to increase their pain.

Mutual reassurance can take the form of acknowledgments, symbolic gestures, or confidence-building measures. To be maximally effective, such steps need to address the other party's central needs and fears as directly as possible. When Sadat spoke to the Israeli Knesset during his dramatic visit to Jerusalem in 1977, he clearly acknowledged Egypt's past hostility toward Israel and, thus, validated Israelis' own experiences. In so doing, he greatly enhanced the credibility of the change in course that he was announcing. At the opening of this visit, Sadat's symbolic gesture of engaging in a round of cordial handshakes with the Israeli officials who had come to greet him broke a long-standing taboo. By signaling the beginning of a new relationship, it had an electrifying effect on the Israeli public. In deep-rooted conflicts, acknowledgment of what was heretofore denied—in the form of recognition of each other's humanity, nationhood, rights, grievances, and interpretation of history—is an important source of reassurance that the other party may indeed be ready to negotiate an agreement that addresses one's own fundamental concerns. By signaling acceptance of the other side's legitimacy, each party reassures the other that negotiations and concessions no longer constitute mortal threats to its security and national existence. By confirming the other side's narrative, each party reassures the other that a compromise does not represent an abandonment of its identity.

An influence strategy based on responsiveness to each other's needs and fears, and the resulting search for ways in which to reassure and benefit each other, has important advantages from a long-term point of view. It does not merely elicit specific desired behaviors from the other party but can contribute to a creative redefinition of the conflict, joint discovery of mutually satisfactory solutions, and transformation of the relationship between the parties.

2.4. Conflict as an Interactive Process with an Escalatory, Self-Perpetuating Dynamic

The influence strategies employed in a conflict relationship take on special significance in light of the fourth proposition, which views international conflict as an interactive process and not merely a sequence of action and reaction by stable actors. In intense conflict relationships, the natural course of interaction between the parties tends to reinforce and deepen the conflict rather than reduce and resolve it. The interaction is governed by a set of norms and is guided by a set of images that create an escalatory, self-perpetuating dynamic. This dynamic can be reversed through skillful diplomacy, imaginative leadership, third-party intervention, and institutionalized mechanisms for managing and resolving conflict. But in the absence of such deliberate efforts, the spontaneous interaction between the parties is more likely than not to increase distrust, hostility, and the sense of grievance.

The needs and fears of parties engaged in intense conflict impose perceptual and cognitive constraints on their processing of new information, with the resulting tendency to underestimate the occurrence and the possibility of change. The ability to take the role of the other is severely impaired. Dehumanization of the enemy makes it even more difficult to acknowledge and access the perspective of the other party. The inaccessibility of the other's perspective contributes significantly to some of the psychological barriers to conflict resolution described by Ross and Ward in a 1995 review. The dynamics of conflict interaction tend

to entrench the parties in their own perspectives on history and justice. Conflicting parties display particularly strong tendencies to find evidence that confirms their negative images of each other and to resist evidence that would seem to disconfirm these images. Thus, interaction not only fails to contribute to a revision of the enemy image but actually helps to reinforce and perpetuate it. Interaction guided by mirror images of a demonic enemy and a virtuous self (a concept introduced by Bronfenbrenner in 1961 and elaborated by White in 1965 and elsewhere) creates self-fulfilling prophecies by inducing the parties to engage in the hostile actions they expect from one another.

Self-fulfilling prophecies are also generated by the conflict norms that typically govern the interaction between parties engaged in an intense conflict. Expressions of hostility and distrust toward the enemy are not just spontaneous manifestations of the conflict; they are normatively prescribed behaviors. The assumption of political leaders that their publics' evaluations of them depends on their adherence to these norms influences their tactical and strategic decisions, their approach to negotiations, their public pronouncements, and (ultimately) the ways in which they educate their own publics. For the publics, in turn, adherence to these norms is often taken as an indicator of group loyalty. Thus, the discourse in deep-rooted conflicts is marked by mutual delegitimization and dehumanization. Interaction governed by this set of norms—at both the microlevel and macrolevel—contributes to escalation and perpetuation of the conflict. Parties that systematically treat each other with hostility and distrust are likely to become increasingly hostile and untrustworthy.

The dynamics of conflict interaction create a high probability that opportunities for conflict resolution will be missed. Parties whose interaction is shaped by the norms and images rooted in the history of the conflict are systematically constrained in their capacity to respond to the occurrence and possibility of change. Each party finds it difficult to communicate the changes that have occurred on its own side, to notice the changes that have occurred on the other side, and to explore the possibilities for change that would serve both sides' interests. Therefore, conflict resolution efforts require promotion of a different kind of interaction that is capable of reversing the escalatory and self-perpetuating dynamics of conflict—an interaction conducive to sharing perspectives, differentiating the enemy image, and developing a language of mutual reassurance and a new discourse based on the norms of responsiveness and reciprocity.

3. PSYCHOLOGICAL PROCESSES PROMOTING CONFLICT

Social–psychological analysis can be particularly helpful in explaining why and how, once a conflict has started, normative and perceptual processes that promote its escalation and perpetuation are set into motion and create or intensify barriers to conflict resolution. By the same token, social–psychological analysis, in helping to identify and understand these barriers, can also suggest ways in which to overcome them.

3.1. Normative Processes

A variety of interaction processes at the mass and elite levels of conflicting societies that influence the evolving course of the conflict are governed by a set of powerful social norms that tend to encourage actions and attitudes conducive to the generation, escalation, and perpetuation of conflict and that tend to inhibit the perception and occurrence of change in the direction of tension reduction and conflict resolution.

3.1.1. Formation of Collective Moods

With periodic shifts in collective mood, public opinion can act as both a resource and a constraint for political leaders in the foreign policy process. In principle, public opinion can provide support for either aggressive or conciliatory policies, but under the prevailing norms in an intense protracted conflict, leaders are more likely to expect—and mobilize—public support for aggressive policies than for conciliatory policies. Apart from transitory moods, certain pervasive states of consciousness underlie public opinion in a society engulfed in a deep-rooted conflict, reflecting the existential concerns and central national narratives widely shared within the population. In many cases, such as Serbia, Northern Ireland, and the Middle East, historical traumas serve as the points of reference for current events. Although these memories may be manipulated by demagogic leaders, they are part of the people's consciousness and are available for manipulation, as is the associated sense of injustice, abandonment, and vulnerability. The effect of such collective moods is to bring to the fore powerful social norms that support escalatory actions and inhibit moves toward compromise and accommodation. When fundamental concerns about survival and identity are tapped, national leaders, with full expectation of public support, are far

more ready to risk war than to take risks for peace, in line with the proposition derived from prospect theory (as applied to international relations by Levy and others) that people are more reluctant to take risks to achieve gains than to avoid losses. Any change in the established view of the enemy and of the imperatives of national defense comes to be seen as a threat to the nation's very existence.

3.1.2. Mobilization of Group Loyalties

Public support is an essential resource for political leaders engaged in a conflict relationship, both in ensuring the public's readiness to accept the costs that their policies may entail and in enhancing the credibility of their threats and promises to the other side. The primary means of gaining public support is the mobilization of group loyalties. Arousal of nationalist and patriotic sentiments, particularly in a context of national security and survival, is a powerful tool in garnering public support. It may evoke automatic endorsement of the policies that the leadership defines as necessary and a willingness to make sacrifices that cannot be entirely understood in terms of rational calculations of costs and benefits. The nation generates such powerful identifications and loyalties because it brings together two central psychological dispositions: the need for self-protection and the need for self-transcendence.

Group loyalties can potentially be mobilized in support of conciliatory policies. Political leaders may promote painful compromises and concessions to the adversary on the grounds that the security, well-being, integrity, and survival of the nation require such actions. Indeed, leaders with impeccable nationalist credentials, such as Charles de Gaulle, Yitzhak Rabin, and F. W. de Klerk, are often most effective in leading their populations toward peaceful resolutions of conflicts once they have decided that this approach best serves their national interests. In general, however, group loyalties are more readily available to mobilize support for aggressive policies than for conciliatory ones. Proposals for aggressive actions can more easily rely on the vocabulary of nationalism, which characteristically marks off the in-group from the out-group to the detriment of the latter. An appeal to defend the nation against an imminent attack, in particular, is more compelling than an appeal to seize a promising opportunity, as prospect theory might predict. Also, such an appeal elicits a nearly unanimous response among members of the population, whereas an appeal to take advantage of an opportunity for peace holds no attraction to that segment of the population that equates peace with surrender.

Processes of group loyalty create barriers to change in a conflict relationship. Group loyalty requires adherence to the group's norms, which call for a militant, unyielding, and suspicious attitude toward the enemy in an intense conflict. Militancy and intransigence thus become the measures of loyalty. Hence, particularly in situations of perceived national crisis, the militants exercise disproportionate power and often a veto over official actions and policies. They impose severe constraints on the ability of leaders to explore peaceful options. Dissent from the dominant conflict norms becomes defined as an act of disloyalty and is suppressed, further undermining the exploration of peaceful alternatives.

3.1.3. Decision-Making Processes

The way in which decisions are made in a conflict situation tends to inhibit the search for alternatives and the exploration of new possibilities, particularly when decision makers are operating in an atmosphere of crisis. These tendencies are by no means inevitable, and there are historical instances of creative decision making in dangerous crisis situations, for example, the Cuban missile crisis (as analyzed by Allison in 1971). However, conflict norms do impose serious burdens on the decision-making process.

A major source of reluctance to explore new options is the domestic constraints under which decision makers labor. In an intense conflict situation, adherence to the conflict norms tends to be seen as the safest course of action. Cautious decision makers assume that they are less vulnerable domestically if they stay with the conflict's status quo, adhere to a discourse of hostility and distrust vis-à-vis the other side, and/or threaten escalatory actions than if they take steps toward accommodation and compromise. The search for alternatives in response to changing realities is also inhibited by institutionalized rigidities in the decision-making apparatus. Decision makers and their bureaucracies operate within a framework of assumptions about available choices, effective strategies, and constituency expectations—shaped by the prevailing conflict norms—that may make them unaware of the occurrence and possibility of change. Furthermore, they often rely on established procedures and technologies, which are more likely to be geared toward pursuing the conflict—by military and other means—than toward resolving it.

The microprocesses of action and interaction in crisis decision making further inhibit the exploration of new options. At the level of individual decision makers, the stress they experience in situations of crisis, when consequential decisions must be made under severe time pressures, limits the number of alternatives they consider and impels them to settle quickly on the dominant response. In intense conflicts, this dominant response is likely to be aggressive and escalatory, as shown in books by Holsti in 1972 and Lebow in 1987. At the level of decision-making groups, crisis decision making often leads to what Janis termed "groupthink," that is, a concurrence-seeking tendency designed to maintain the cohesiveness of the groups. Decision making under these circumstances is much more likely to produce policies and actions that perpetuate and escalate the conflict than to produce innovative ideas for conflict resolution.

3.1.4. Negotiation and Bargaining Processes

The norms governing negotiation and bargaining between parties involved in long-standing conflict strongly encourage zero-sum thinking, which equates the enemy's loss with one's own gain. Negotiation, even distributive bargaining in its narrowest form, is possible only when the parties define the situation—at least at some level—as a win–win, mixed-motive game in which they have both competitive and cooperative goals. Each party, while pursuing its own interests, must actively seek out ways in which the adversary can also win and appear to be winning. But this is precisely the kind of effort that is discouraged by the conflict norms.

At the microlevel, negotiators in an intense conflict tend to evaluate their performance by the forcefulness with which they present their own case and by their effectiveness in resisting compromise. To listen to what the other side needs and to help the other side achieve its goals would violate the conflict norms and might subject the negotiators to criticism from their own constituencies, particularly from their hard-line domestic opposition. At the macrolevel, the parties tend to pursue an overall outcome that strengthens their own strategic position and weakens that of the adversary, even when they recognize their common interest in negotiating certain specific issues. Such a strategy reduces the other party's incentive for concluding an agreement and its ability to mobilize public support for whatever agreement is negotiated. Zero-sum thinking at both levels undermines the negotiating process, causing delays, setbacks, and repeated failures.

3.1.5. Structural and Psychological Commitments

Finally, conflict creates certain commitments that take on a life of their own and contribute to structural changes conducive to the escalation and perpetuation of the conflict, as described in Rubin and colleagues' *Social Conflict* in 1994. Most obviously, in a conflict of long standing, various individuals, groups, and organizations (e.g., military, political, industrial, scholarly) develop a vested interest in maintaining the conflict as a source of profit, power, status, and/or raison d'être. Others, although not benefiting from the conflict as such, may have a strong interest in forestalling a compromise solution because it would not address their particular grievances or fulfill their particular aspirations. Vested interests do not necessarily manifest themselves in deliberate attempts to undermine efforts at conflict resolution. They may take indirect and subtle forms such as interpreting ambiguous realities and choosing between uncertain policy alternatives in ways that favor continuation of the conflict.

Vested interests and similar structural commitments to the conflict are bolstered by psychological commitments. People involved in a long-standing and deep-rooted conflict tend to develop a worldview that is built around the conflict and would be threatened by an end to the conflict. Resistance to change is likely to be more pronounced when the cognitive structure or ideology in which the view of the conflict is embedded is more elaborate, because changing this view would have wider ramifications. In an intense conflict, the image of the enemy is often a particularly important part of people's worldview, with implications for their national identity, view of their own society, and interpretation of history. This is one reason why images of the enemy are highly resistant to change and contribute to the escalatory and self-perpetuating dynamic of conflict.

3.2. Perceptual Processes

Perceptual and cognitive processes, that is, the ways in which people interpret and organize conflict-related information, play a major role in the escalation and perpetuation of conflict and create barriers to redefining and resolving the conflict despite changing realities and interests. Two perceptual processes that characterize mutual images of parties in conflict can account for this effect: the formation of mirror images and the resistance of images to contradictory information.

3.2.1. Mirror Image Formation

As noted earlier, Bronfenbrenner and White, social psychologists writing about U.S.–Soviet relations, first noted the formation of mirror images as a characteristic of many conflict relationships. Both parties tend to develop parallel images of the self and the other, except with the values reversed. The core content of mirror images is captured by the good–bad dimension. Each side sees itself as virtuous and peaceful, arming only for defensive reasons and prepared to compromise. The enemy, in contrast, is seen as evil and hostile, arming for aggressive reasons and responsive only to the language of force.

A typical corollary of the good–bad images in protracted conflicts is the view that the other party's aggressiveness is inherent in its nature (e.g., ideology, religion, national character, political system), whereas any signs of aggressiveness on one's own part are entirely reactive and defensive. In the language of attribution theory, the enemy's aggression is explained in dispositional terms, whereas one's own aggression is explained in situational terms. Another common corollary of the good–bad image, one that derives from the virtuous self-image, is the assumption on each side that the enemy knows very well that "we" are not threatening them. Each side's own basic decency and peacefulness, and the provocation to which it has been subjected, are so obvious to one's own party that they must also be obvious to the other side. Apart from such generic features of mirror images, which arise from the dynamics of intergroup conflict across the board, mirror images in any given case may reflect the dynamics of the specific conflict. Thus, ethnic conflicts may be characterized by mutual denial of the other side's national identity accompanied by efforts to delegitimize the other's national movement and claim to nationhood, by mutual fear of national and personal annihilation, by a mutual sense of victimization by the other side, and/or by a mutual view of the other side as a source of one's own humiliation and vulnerability.

The mirror image concept implies that certain symmetries in the parties' reactions arise from the very nature of conflict interaction and that they play an important role in escalating the conflict. There is no assumption that all images of the self and the enemy are mirror images, that images on the two sides are equally inaccurate, or that there is empirical symmetry in the two sides' historical experiences and current situation or moral equivalence in their positions. However, the dynamics of the conflict relationship produce a degree of parallelism in some of the images developed by both participants in that relationship, arising out of the motivational and cognitive contexts in which they operate. At the level of motivation, each side is concerned with "looking good" when blame for the conflict events is being apportioned. Therefore, political leaders feel a strong need to persuade themselves, their own people, the rest of the world, and future historians that the blame rests with the enemy. Cognitively, each side views the conflict from its own perspective and is convinced that it is acting defensively and with the best intentions (painfully aware of its own needs, fears, historical traumas, grievances, suspicions, and political constraints) and that this is so self-evident that it must be equally clear to the enemy.

Mirror images produce a spiraling effect (exemplified by the classical pattern of an arms race) because each side interprets any hostile action by the other as an indication of aggressive intent against which it must defend itself, yet that side's own reactions—the defensive nature of which presumably should be obvious to the enemy—are taken by the other side as signs of aggressive intent. The effect of mirror images is accentuated insofar as the enemy's ideology or national character is perceived to be inherently aggressive and expansionist. In addition to the escalatory effect of mirror images, they tend to make conflicts more intractable because the sharp contrast between the innocent self and the aggressive other makes it difficult to break out of a zero-sum conception of the conflict. However, the concept of mirror images may be a useful tool in conflict resolution. For example, in problem-solving workshops, the parties' discovery that their own actions are perceived differently by the other side than by themselves may open them up to the possibility that the reverse may be true as well. Thus, they may gain access to each other's perspective, insight into the escalatory effects of such two-directional differences in perception, and awareness of the need for mutual reassurance to set a deescalatory process in motion.

3.2.2. Resistance to Contradictory Information

The second feature of conflict images, their high propensity to resisting contradictory information, inhibits the perception of change and the expectation of future change. A great deal of social–psychological theorizing and research has addressed the general phenomenon of the persistence of attitudes and beliefs in the face of new information that, from an outside point of view, challenges their validity but is somehow neutralized or

ignored. Research has focused on several types of mechanisms that account for resistance to contradictory information, including selectivity, consistency, attribution, and the self-fulfilling prophecy. The concepts of selective exposure, selective perception, and selective recall all point to the fact that people's attitudes help to determine the kind of information that is available to them. People are more likely to seek out and be exposed to information that confirms their existing attitudes and to perceive and remember new information in ways that fit into their preexisting cognitive framework. The various models of cognitive consistency (e.g., Heider's theory of cognitive balance, Festinger's theory of cognitive dissonance) suggest that, in the interest of maintaining consistency, people tend to screen out information that is incongruent with their existing beliefs and attitudes. Although inconsistent information may also instigate attitude change, it is more likely to be resisted when the existing attitudes are strongly held and have wide ramifications, as is the case with enemy images. Attribution mechanisms, linked to what Ross called "the fundamental attribution error" in 1977, also promote confirmation of the original enemy image. Hostile actions by the enemy tend to be attributed dispositionally, providing further evidence of the enemy's inherently aggressive, implacable character, whereas conciliatory actions are explained away as reactions to situational forces, requiring no revision of the original image (a phenomenon observed in research by Rosenberg and Wolfsfeld in 1977, by Heradstveit in 1981, and by Rouhana in 1997). Finally, interactions between conflicting parties tend to create self-fulfilling prophecies by causing one's adversaries to behave in line with one's expectations—to take on the roles in which they have been cast by the other side—thereby confirming the parties' original attitudes.

The mechanisms that account for resistance to disconfirming information are particularly powerful in a conflict relationship for several reasons. First, images of the enemy and conflict-related self-images are central aspects of the national consensus; therefore, resistance to disconfirming information is reinforced by strong normative pressures. Second, in a conflict relationship, the opportunities and capacity for taking the perspective of the other side are limited, and this reduces the impact of potentially new information about the varieties, changes, and signs of flexibility in the other side's views. Third, the resistance of enemy images to disconfirmation is magnified by strong beliefs about the unchangeability of the enemy and is reinforced by the view that it is dangerous or even treasonous to propose that the enemy has changed or will change.

Despite all of the reasons why conflict images are particularly resistant to contradictory information, they are not immutable. Social–psychological evidence suggests that they can change, and historical evidence shows that they do change. The challenge for scholars and practitioners of international conflict resolution is to devise the means to overcome their resistance to change.

4. IMPLICATIONS FOR CONFLICT RESOLUTION

Social–psychological analysis can contribute to international peacemaking by identifying the psychological and social processes that generate and escalate violent conflicts and impede their peaceful resolution as well as by identifying the conditions and procedures required for breaking and reversing the conflict cycle and setting into motion a process of change in the direction of conflict resolution and reconciliation. Social–psychological principles have also informed the development and application of various unofficial microlevel efforts at conflict resolution, designed to complement official diplomacy in a larger multidimensional peace process. Efforts along these lines were pioneered by Burton, who described and conceptualized them in his 1969 book and in numerous subsequent publications. Fisher's *Interactive Conflict Resolution* in 1997 summarized and integrated the range of psychologically based models for intervening in protracted conflicts between identity groups. Kelman's model, interactive problem solving, is described in a 1998 article and a series of other publications, as are the problem-solving workshops through which the interactive problem-solving model is operationalized.

The implications of social–psychological analysis for the macroprocess of conflict resolution, as well as for the microprocess of problem-solving workshops and related activities, can be summarized by returning to the four propositions about the nature of international conflict spelled out earlier in this article.

The view of conflict as a process driven by collective needs and fears serves as a reminder that a conflict cannot be genuinely resolved until these needs and fears of the parties are addressed. Only solutions that satisfy the fundamental needs and allay the deepest fears of each party can serve as the basis of a stable, long-term peace and a cooperative, mutually enhancing

relationship. Unofficial interactive approaches to conflict resolution, such as problem-solving workshops, can contribute to resolving contentious issues in ways that meet the basic needs of one party without raising the fears of the other. They do so by enabling unofficial but politically involved representatives of the two parties to engage in exploratory noncommittal communication in which they can enter into each other's perspective, gain insight into the other's (and indeed their own) main concerns, and jointly shape creative new ideas for mutually satisfactory solutions to the conflict. The insights and ideas generated by this process can then be injected into the political debate and the decision-making process within the two societies.

The view of conflict as an intersocietal process points to the limits of political agreements signed by governments, often under the pressure, or with the mediation, of outside powers or international organizations. At least in the case of a protracted existential conflict between identity groups, a durable peace requires movement beyond settlement of the conflict to its resolution. The agreement must address both parties' basic concerns, must be widely accepted within the two societies, and must reflect and further promote structural and attitude change conducive to a transformation of the relationship between the societies. Another implication of the intersocietal nature of conflict is a view of diplomacy as a complex array of complementary official and unofficial processes. Unofficial interactions make it possible to explore potential openings for mutual accommodation and help to create a political atmosphere that is conducive to negotiation. Furthermore, awareness of the intersocietal nature of conflict helps to counteract the monolithic image that parties in conflict tend to have of each other by encouraging them to attend to what is happening within the other society and to the diversity of tendencies that it encompasses.

Problem-solving workshops and related forums for interactive conflict resolution can be seen as coalitions across conflict lines, enabling pro-negotiation elements on the two sides to find common ground in the pursuit of their respective interests. They engage in a process of direct interaction designed to generate ideas for resolving the conflict that address the needs of both parties and to which both parties can feel committed—essential conditions for agreements that will be widely accepted within the two societies and, thus, conducive to a durable peace. Binding agreements can be reached only at the official level, but unofficial interactions can contribute to enhancing the probability and quality of official agreements. They can also promote and model the new relationship between the conflicting societies on which a long-term peace must ultimately rest.

The third proposition, positing conflict as a multifaceted process of mutual influence, suggests which strategies and tactics of influence between conflicting parties are most conducive to conflict resolution and to the development of a long-term peaceful relationship. Coercive strategies, which are the dominant mode of influence in conflict relationships, entail serious costs and risks and tend to produce only short-term effects. Therefore, there is a need for a shift of emphasis from the use or threat of force to the use of positive incentives. Moreover, positive inducements are most effective when they are deliberately chosen to be responsive to the needs, fears, and constraints of the other party. A key element of an influence strategy based on responsiveness is mutual reassurance, which can take the form of acknowledgments, symbolic gestures, and/or confidence-building measures. A strategy of responsiveness and reassurance is most effective when it addresses the other party's most basic concerns about identity and security and when it adheres to the principle of reciprocity.

The unofficial, noncommittal setting of problem-solving workshops can be particularly useful in identifying and formulating acknowledgments, gestures, and confidence-building measures that each side can offer to the other. Through a process of joint thinking, the parties can help each other to devise steps of mutual reassurance that would be meaningful to the recipient without entailing unacceptable costs to the actor. This process often involves some informal negotiation, including "negotiation" of national identities and national narratives. Parties cannot be expected to abandon their identities or their national narratives, but each party may be able to adjust its own identity so that it is no longer contingent on denial of the other party's identity and to accommodate its own narrative to the narrative of the other.

Finally, the view of conflict as an interactive process with an escalatory, self-perpetuating dynamic suggests that conflict resolution efforts must be designed to counteract and reverse this conflict dynamic. A conflict relationship generates images and norms that entrench the conflict and create barriers to change that inhibit conflict resolution. Therefore, conflict resolution efforts must be geared to discovering the possibilities for change, identifying the conditions for change, and overcoming the resistances to change. Openness to change and reversal of the conflict dynamic depend on the

establishment of a new discourse among the parties characterized by a shift in emphasis from power politics and threat of coercion to mutual responsiveness, reciprocity, and invitation to a new relationship.

At the microlevel, problem-solving workshops and similar approaches to conflict resolution are designed to contribute to the reversal of the escalatory dynamic of conflict interaction by promoting a different kind of interaction characterized by a deescalatory dynamic. To this end, workshops encourage the parties to penetrate each other's perspective, to differentiate their images of the enemy, to develop a deescalatory language and ideas for mutual reassurance, and to engage in joint problem solving designed to generate ideas for resolving the conflict that are responsive to the fundamental needs and fears of both sides.

See Also the Following Articles

Decision Making ■ Groupthink ■ Human Rights

Further Reading

Allison, G. T. (1971). *Essence of decision: Explaining the Cuban missile crisis*. Boston: Little, Brown.

Bronfenbrenner, U. (1961). The mirror image in Soviet–American relations: A social psychologist's report. *Journal of Social Issues*, *17*(3), 45–56.

Burton, J. W. (1969). *Conflict and communication: The use of controlled communication in international relations*. London: Macmillan.

Burton, J. W. (Ed.). (1990). *Conflict: Human needs theory*. New York: St. Martin's.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Fisher, R. J. (1997). *Interactive conflict resolution*. Syracuse, NY: Syracuse University Press.

George, A. L., & Smoke, R. (1974). *Deterrence in American foreign policy: Theory and practice*. New York: Columbia University Press.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: John Wiley.

Heradstveit, D. (1981). *The Arab–Israeli conflict: Psychological obstacles to peace* (2nd ed.). Oslo, Norway: Universitetsforlaget.

Holsti, O. R. (1972). *Crisis, escalation, war*. Montreal, Canada: McGill–Queen's University Press.

Janis, I. L. (1982). *Groupthink* (2nd ed.). Boston: Houghton Mifflin.

Jervis, R. (1976). *Perceptions and misperceptions in international politics*. Princeton, NJ: Princeton University Press.

Jervis, R., Lebow, R. N., & Stein, J. G. (1985). *Psychology and deterrence*. Baltimore, MD: Johns Hopkins University Press.

Kelman, H. C. (Ed.). (1965). *International behavior: A social–psychological analysis*. New York: Holt, Rinehart & Winston.

Kelman, H. C. (1991). A behavioral–science perspective on the study of war and peace. In R. Jessor (Ed.), *Perspectives on behavioral science: The Colorado Lectures* (pp. 245–275). Boulder, CO: Westview.

Kelman, H. C. (1993). Coalitions across conflict lines: The interplay of conflicts within and between the Israeli and Palestinian communities. In S. Worchel, & J. Simpson (Eds.), *Conflict between people and groups* (pp. 236–258). Chicago: Nelson–Hall.

Kelman, H. C. (1998). Social–psychological contributions to peacemaking and peacebuilding in the Middle East. *Applied Psychology: An International Review*, *47*, 5–28.

Kelman, H. C., & Bloom, A. H. (1973). Assumptive frameworks in international politics. In J. N. Knutson (Ed.), *Handbook of political psychology* (pp. 261–295). San Francisco: Jossey–Bass.

Kelman, H. C., & Fisher, R. J. (2003). Conflict analysis and resolution. In D. O. Sears, L. Huddy, & R. Jervis (Eds.), *Oxford handbook of political psychology* (pp. 315–353). New York: Oxford University Press.

Lebow, R. N. (1987). *Nuclear crisis management: A dangerous illusion*. Ithaca, NY: Cornell University Press.

Levy, J. S. (1992). Prospect theory and international relations: Theoretical applications and analytical problems. *Political Psychology*, *13*, 283–310.

Osgood, C. E. (1962). *An alternative to war or surrender*. Urbana: University of Illinois Press.

Rosenberg, S. W., & Wolfsfeld, G. (1977). International conflict and the problem of attribution. *Journal of Conflict Resolution*, *21*, 75–103.

Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 173–220). New York: Academic Press.

Ross, L., & Ward, A. (1995). Psychological barriers to dispute resolution. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 27, pp. 255–304). San Diego, CA: Academic Press.

Rouhana, N. N. (1997). *Palestinian citizens in an ethnic Jewish state: Identities in conflict*. New Haven, CT: Yale University Press.

Rubin, J. Z., Pruitt, D. G., & Kim, S. H. (1994). *Social conflict: Escalation, stalemate, and settlement* (2nd ed.). New York: McGraw–Hill.

Schelling, T. C. (1963). *The strategy of conflict*. Cambridge, MA: Harvard University Press.

Tetlock, P. E., Husbands, J. L., Jervis, R., Stern, P. C., & Tilly, C. (Eds.). (1989). *Behavior, society, and nuclear war* (Vol. 1). New York: Oxford University Press.

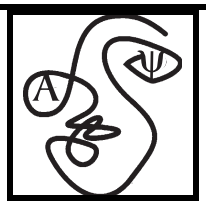
Tetlock, P. E., Husbands, J. L., Jervis, R., Stern, P. C., & Tilly, C. (Eds.). (1991). *Behavior, society, and nuclear war* (Vol. 2). New York: Oxford University Press.

White, R. K. (1965). Images in the context of international conflict: Soviet perceptions of the U.S. and the U.S.S.R. In H. C. Kelman (Ed.), *International behavior: A social–psychological analysis* (pp. 238–276). New York: Holt, Rinehart & Winston.

White, R. K. (1984). *Fearful warriors: A psychological profile of U.S.–Soviet relations*. New York: Free Press.

Zartman, I. W. (1989). *Ripe for resolution* (2nd ed.). New York: Oxford University Press.

Zartman, I. W., & Rasmussen, J. L. (Eds.). (1997). *Peacemaking in international conflict: Methods and techniques*. Washington, DC: U.S. Institute of Peace Press.



Internet Counseling

Azy Barak
University of Haifa, Haifa, Israel

1. Introduction
2. Historical Development
3. The Practice of Internet Counseling
4. Research on Internet Counseling
5. Ethical, Professional, and Legal Considerations
6. Limitations and Problems
7. Summary
Further Reading

GLOSSARY

blog (Web log) A publicly accessible Web page, updated regularly and often, that serves as a personal journal of the site owner.

chat Text-based synchronous (real-time) conversation between two Internet users through special software.

chat room A virtual environment in which numerous Internet users may "meet" and converse with one another in synchronous communication through typed text.

cyberspace The virtual space that people enter and experience when connected to the Internet, following an idea used in Gibson's *Neuromancer*.

e-mail Abbreviated term for electronic mail; a typed letter sent to a virtual mailbox over the Internet.

e-mail list (listserv) Electronic mailing list where e-mails are sent by subscribers to all other subscribers on the list.

emoticon A cluster of punctuation or a small graphical sign or picture that may be used in online communication to compensate for the lack of nonverbal cues to indicate a facial expression or gesture.

encryption Coding of text through special software so that only the intended receiver of the message can decipher it (i.e., decryption).

forum (message board or discussion board) A Web site designated to allow asynchronous group communication.

instant messaging Software that allows text-based synchronous communication, with possible time delays (i.e., users control level of synchronicity), between two Internet users.

post Publishing a message on the Internet, including sending a message to an e-mail list or publishing a message in a forum.

synchronous Coinciding in time of communication that is carried out with all parties present at the same time (e.g., chat).

thread A multiple-part virtual "conversation" on a given topic in an e-mail list or a forum composed by a leading (first) message and responses.

virtual communication Communication between two or more anonymous individuals through the mediation of computers where the looks, appearance, and nature of these individuals are basically the products of their imaginations.

Internet counseling is a general term referring to a wide spectrum of approaches and techniques that exploit the Internet's special capabilities in facilitating counseling and psychotherapy online. Growing evidence suggests that despite the lack of in-person contact and the consequent lack of nonverbal cues, professionals do provide effective interventions online. However, in addition to focused training in Internet counseling,

©2004 Elsevier Inc.
All rights reserved.

special precautions are necessary, especially regarding unique ethical and legal issues.

1. INTRODUCTION

Internet counseling is a generic term that refers to various methods used by psychologists to provide professional services through, or with the help of, the Internet. Common terms that have been used include online (psycho)therapy (or counseling), e-therapy, cybertherapy, and virtual therapy. Although attempts have been made to differentiate among these terms, it has become clear that they refer to similar, if not identical, processes and techniques. Internet counseling involves a variety of online communication technologies that generally fit into 2×2 dimensions: synchronous versus asynchronous communication and individual versus group communication. Individual synchronous communication is conducted by means of personal chat or instant messaging, individual asynchronous communication is conducted by e-mail exchanges, group synchronous communication is carried out though a chat room, and group asynchronous communication is managed by a forum (discussion board) or an e-mail list (listserv). Internet counseling may also include other online methods and tools such as Web site-accompanied therapy and online personal journals. Although currently less prevalent, Internet counseling may also use computer sound and camera (Webcam or videoconferencing) facilities.

Internet counseling is usually text based. Consequently, it is characterized by the invisibility of counselors and clients, lack of nonverbal and physical cues, and absence of eye contact—three factors that significantly influence social communication and interpersonal relations. These characteristics, in addition to other factors such as (possible) anonymity and aloneness by the computer, create the "online disinhibition effect" that brings about greater and faster openness and closeness, on the one hand, and less controlled behavioral expressions (often aggression), on the other. Thus, the therapeutic process accelerates more intensively and rapidly in comparison with face-to-face therapy, resulting in faster disclosures and greater exposure of intimate contents. In addition, unique features of Internet use—availability, flexibility of use (in terms of time and place), affordability, anonymity, acceptability, aloneness, interactivity, hypertextuality, updatability, multimedia capability, easy departure (escape), and richness of contents—have gradually caused Internet counseling to become a convenient accepted channel through which people seek psychological consultation and help.

2. HISTORICAL DEVELOPMENT

Although not documented, it seems that serious professional attempts to harness the Internet for therapeutic purposes were first made during the mid-1990s. At that time, e-mail was used as the primary technology to communicate between practitioners and clients, either individually or in groups. Despite positive experiences reported by therapists, traditional mental health practitioners voiced their opposition to the new method, citing its limitations and possible risks. Five major problems were highlighted: (a) counselors' inability to accurately and correctly perceive and evaluate clients' full emotional states due to the lack of visual and physical cues (e.g., tears, smiles, gestures, body movements); (b) therapists' difficulty in communicating warmth and empathy owing to a similar lack of nonverbal cues; (c) the unclear, intangible, easy-to-break boundaries between therapists and patients; (d) the possibility of erroneously dealing with clients from other countries and cultures without being aware of indigenous customs, rules, norms, or other pertinent information; and (e) the risk of dealing with severe pathological cases, especially suicidal clients, brought on by problems relating to identifying and locating them at critical moments. Criticisms also were made in regard to problems of privacy, confidentiality, and security; the possible undesirable effects of clients' individual disadvantages, such as those related to computer and verbal skills, on the therapeutic process and outcome; discrimination against economically distressed potential clients (what has been termed "the digital divide"); and the potential for impersonation, lying, and other abuses of the system. Several professional associations, in addition to individual professionals, published warnings to prevent (or minimize) online counseling or at least to try restricting its possible dangers to clients.

However, with the help of continuously upgraded hardware and software, as well as growing awareness of and experience, training, and pioneering research in this developing field, many of the criticisms have been answered and gradually been overcome. Moreover, arguments began to be advanced that the opposition reflected mainly resistance to change in general and ignored (or downplayed) a growing body of evidence that supported and pointed to the success of Internet counseling.

Evidently, Internet counseling has now been established and is generally accepted as legitimate professional conduct. During recent years, a growing number of publications—theoretically, practically, and empirically oriented—have emerged in professional outlets as well as in professional conferences. A major portal (www.metanoia.org) lists many relevant issues relating to psychotherapy through the Internet. Quite a few professional and scientific books have been published on the subject. A major professional association, the International Society for Mental Health Online, has been formed in an attempt to provide a common organization for professionals interested in this emerging area and to promote the understanding, use, and development of mental health online.

3. THE PRACTICE OF INTERNET COUNSELING

Internet counseling is diverse, heterogeneous, and multifaceted. The following 10 points describe its characteristics and provide some distinctions from traditional counseling. First, unlike traditional face-to-face therapy, Internet counseling can be practiced through various alternative communication channels. These channels differ from one another in terms of synchronicity (usually synchronous or asynchronous communication, whereas instant messaging allows a changing degree of synchronicity), mode (individual or group communication), means of communication (usually typed text, but pictures and voice communication are also possible), and the degree of on-site human involvement (direct human interaction, self-help by using information and instructions published on a Web site, or online interactive software). Thus, clients have a wide range of choices. Counselors may specialize and offer a specific service modality and provide it through numerous channels. Likewise, clients may receive service through one or more channels of communication.

Second, Internet counseling may be practiced as a complementary, adjacent, or adjunctive process to face-to-face counseling or as an independent process of therapy. As an accompanying procedure, clients may use online communication with counselors in between their face-to-face sessions (according to preestablished agreed-on rules); manage online diaries (i.e., specific behavior monitoring and recording) relevant to the therapeutic process that may be inspected and reviewed by themselves and their counselors at a time and place of choice; deliver materials relevant to the therapy process such as articles, important links, and necessary or urgently needed documents; and publish and maintain a Web site in which consequences or the offspring thoughts of face-to-face sessions may be expressed or implemented (i.e., a blog). Cyberspace may also be used to search for materials and information found to be relevant in the face-to-face meetings or for exercises and practice activities suggested in face-to-face sessions. In all of these procedures, counselors attempt to integrate offline and online therapeutic experiences at different rates and to different degrees to promote clients' change. Independently, clients and counselors may meet online (usually, but not necessarily, through a designated Web site) and interact virtually through an online channel(s) of communication. In using this procedure for therapy, counselors and clients may be distant from each other geographically, interact in different time zones, and become involved with the therapeutic interaction at a time of choice (if asynchronous communication is used). To overcome the lack of standard, normal, and familiar therapeutic procedures in these cases, such as checking and evaluating therapists' credentials, payment procedures (often done through third-party payment), emergency contacts, and termination of the therapy, a set of instructions and rules is usually established and agreed on between counselors and clients at a preliminary stage. Internet counseling methods, rates of sessions, and lengths of interactions vary among counselors according to their approaches, their availability, the clients' particular problems, communication channels, costs, and other professional considerations.

Third, Internet counseling generally relies on text-based communication between counselors and clients. Writing can often be healing. For many people, venting their personal concerns, expressing their deepest feelings, and sharing their intimate thoughts—all while alone and writing as though talking to themselves, free of therapists' scrutinizing eyes, feeling more secure and protected in their own habitats than in counselors' offices, having the ability to preserve a full record of counseling dialogues, and experiencing less intimidation in developing relationships with virtual (but trustworthy) figures—contribute to creating a unique therapeutic experience, much different from traditional "talking and listening" counseling. The ability of clients to read back, cite, and quote themselves or counselors adds a special value to the counseling dialogue. For many people, sharing their writings—even from the past—with trusted individuals gives them a feeling of relief and sometimes even of elation.

Fourth, Internet counseling may be provided to individuals for either long-term or short-term interventions as well as for single consultations. Although continuous counseling process enables the establishment of relationships between counselors and clients as well as the development of trust and greater exposure, a one-time question-and-answer psychological service is quite common on the Internet. Similar to advice columns in newspapers and magazines, clients present (usually either through sending e-mails or posting messages on online bulletin boards) personal issues and expect counselors to provide advice or solutions. However, there are several differences. First, online consultations are not necessarily public given that clients may discretely send e-mails to counselors and receive private answers from counselors who offer such a service (through Web sites or other means of communication). Second, accumulated questions and answers can be (and actually are) edited and posted on the Internet, classified by problem area, target population, and the like, with the help of an index and an internal search engine for the benefit of many visitors. Third, such an online service is convenient and relatively fast, and it enables people in distress to consult anonymously with experts during times of need for quick assistance.

Fifth, Internet counseling enables consultation and treatment through various theoretical therapeutic approaches, methods, and techniques. For example, psychodynamic, cognitive, and behavioral approaches have been used online, with each one finding the Internet to be a convenient—or even an ideal—platform for interacting with clients according to counselors' own conceptions and principles. For psychodynamic counselors (and even more so for psychoanalytic counselors), cyberspace represents a ready desirable environment for psychotherapeutic encounters because it is characterized by ambiguity. Thus, dynamic processes activated by projection, transference, and object relations usually prevail. In other words, the perceived disadvantages of Internet relationships, such as invisibility and blurred boundaries, become advantageous to counselors who hold to the psychodynamic approach. Therefore, analyzing clients' personality dynamics becomes easier, generally faster, and perhaps more valid than traditional therapy. For cognitive counselors, the straightforward use of clients' verbalization of thoughts and feelings obviously represents a major advantage. Compared with traditional therapy, cognitively oriented counselors are more able to assess and identify thought patterns that may be responsible for negative emotions and consequent problematic behaviors. Because all online verbal interactions with clients are recorded (when using the common text-based therapy), clients' statements are easy to cite, compare, analyze, and manipulate. In this context, it is important to mention that therapists who apply bibliotherapy and therapists who claim that writing and reading can heal also find online, text-based therapy to be a highly effective therapeutic means. Likewise, users of narrative psychotherapy can exploit written interactions with patients to effectively help them construct and reconstruct personal "stories." Yet behaviorally focused counselors also find Internet counseling to be ideal medium in many senses in that it allows a structured, programmed, straightforward, easy to construct and maintain set of instructions to be communicated to clients, whether through personal contact or through Web sites. Moreover, clients' continuous recording of behavioral data and their delivery to counselors are simple and efficient. Counselors may guide, supervise, and monitor clients' behavior management from a distance and use in vivo simulations of various behavior tasks, accompanied by immediate feedback. Various theoretical approaches, then, can be applied through Internet counseling, as can any eclectic combination desired by counselors. It is important to note, however, that the unique psychological environment created in cyberspace might eventually result in a unique or new adaptation of traditional therapeutic approaches that previously were used solely in the face-to-face physical environment.

Sixth, not only does Internet counseling allow more than a single, standard, and invariable way of communication, its alternative modalities also enable a variety of interactions and activities that may enrich the counseling experience and contribute to client change. In addition to the obvious advantage of Internet counseling in terms of flexibility of location and time, several other features provide added value to the interactions. Asynchronous personal communication through e-mail enables the occurrence of several psychological processes that may improve the counseling experience. For clients, the ability to express feelings and thoughts at times of their choosing—especially at times of personal distress—and not at periodic predetermined meetings with counselors, is a significant advantage. The ability to reread counselors' messages as often as desired, and at any time, is also important. Patients' ability to reflect, associate, consider, recall, and process these interactions in between e-mail exchanges contributes to psychologically desirable incubation and reflection processes. Moreover, the ability to go back and compare notes—the clients' own notes or their counselors'

notes—is very useful during the therapy process. In addition, the simplicity of attaching other documents to e-mail messages, including drawings and links, possibly makes e-mail interaction more effective than mere talk. For counselors, e-mail interaction enables referring to clients at times of their choosing and convenience, providing more opportunities to analyze, compare, recall (i.e., retrieve previous messages), use quotes, and enrich response messages with important relevant links and documents. Likewise, counseling through asynchronous group communication—by using e-mail lists or forums—enables all group participants to benefit from the same advantages. An additional special ingredient that exists in asynchronous online groups is the ease of members' interacting individually with one another through a back channel of communication, a factor that might contribute to groups' dynamics and change so long as there are certain rules that participants must follow. Synchronous individual counseling, executed through a chat or an instant messaging program (e.g., ICQ), enables real-time communication between patients and their therapists from nearly any location. Unlike telephone conversations under similar circumstances, the use of computers allows conversations to be saved for future recall, analysis, and/or quoting. The disinhibition effect of online communication brings about a rapid exposure of significant and essential material for counseling. In many instances, clients open up in minutes about matters that would have taken weeks or even months to reveal in face-to-face interaction with therapists. Moreover, by experiencing less intimidation (and, consequently, being less defensive) and feeling more egalitarian and equal in status, patients have a greater chance of constructively processing and adopting therapists' messages. One of the unique features of synchronous communication that is available to counselors only on the Internet is "multiconversing," that is, conversing privately with several individuals in parallel. Although the technology requires certain technical and cognitive abilities and skills that might be acquired and promoted through training and experience, multiconversing allows a counselor to make use of the time required for one client to read the counselor's message, process it, and type a response to communicate with other clients. Although special concentration and refocusing abilities are obviously required, possible benefits of this activity are clear, especially when providing counseling services to a population through an Internet-based support center. However, synchronous online group therapy (through a chat room) is quite difficult to manage due to the lack of nonverbal communication and the difficulty in controlling participants' behavior. If practiced, it is important to have a strict, well thought out set of rules and a small number of participants (usually not more than seven).

Seventh, the practice of Internet counseling allows nonhuman interventions; that is, providing interventions through psychoeducational Websites, as well as through interactive software. Advanced sophisticated Web sites employ interactive software that communicates with users in an attempt to assess and provide tailored intervention in specific problem areas. This is done through various alternative methods, including interactive conversations with robotic therapists, adaptive and tailored instructions in dealing with a specific problems, meditating and relaxation suggestions, and interactive decision-making assistance (e.g., about a career direction or school major). More recently, attempts have been made to adapt virtual reality software in the field to the Internet, thereby allowing mass use and continuous central upgrades. Therefore, the ability of advanced (and continuously developing) software to deliver professional information and interventions in an efficient, attractive, rich, and friendly way makes this a prevalent and pervasive type of communication with clients. Regardless of the areas of distress or problems, whether quitting smoking, obtaining sexuality education, experiencing posttraumatic stress disorder following a sexual assault, facing the death of a loved one, choosing a major in school, or going through a divorce, a well-designed, thoughtful, updated, interactive, sophisticatedly linked (to other Web resources), and stimulating Web site can provide effective psychological interventions for people in need. As in any process of self-help, users' commitment, on the one hand, and counselors' monitoring, on the other, may be problematic. However, the cost–benefit advantage, in addition to factors mentioned earlier (e.g., anonymity), has turned this into a common and successful approach.

Eighth, free Internet counseling, by means of information, advice, and support, has flourished. Despite the clear problem of potential abuse of the system by nonprofessionals or even criminals (e.g., practicing without a license), the use of the Internet has enabled people in need to reach out for psychological help when, for whatever reasons, they could not receive it in standard ways. People in remote locations, the physically disabled, the elderly, the sick, those with financial limitations, and those who are too shy, introverted, or afraid of exposure (e.g., celebrities and their families) all fall into this category. Online support services for a variety of areas of distress are operated by governments,

organizations (e.g., schools), professional associations, and private people for free. Usually, e-mail consultation is offered to individual referrals. One of the main activities of such free counseling services, however, is the provision of online support groups, in which participation is usually open. This type of Internet counseling, with or without the active participation of trained counselors, is widespread and considered to be successful.

Ninth, Internet counseling is effectively and conveniently used as a vehicle for communicating with clients before beginning and after terminating face-to-face counseling. Information provided through Web sites regarding counseling approaches, counseling methods, the counseling process, specific counselors, costs, insurance coverage, and technicalities equip clients with important and relevant knowledge before starting counseling. Questions and unclear matters are easily dealt with through e-mail. This preparatory stage not only saves time and effort but also helps to build realistic constructive expectations about therapy. After termination, e-mail is conveniently used to contact former clients at any point in time, either for professional follow-up or for research purposes (e.g., questionnaires may be attached). On termination, Internet counseling clients are often urged to make e-mail contact with their counselors when needed. This step provides not only a convenient way in which to communicate but also reassurance for patients departing from close, and sometimes dependent, therapeutic relationships. It should be noted, however, that this option might hinder clients from working thorough successful terminations.

Finally, Internet counseling may include and coincide with online psychological assessment procedures. Online psychological tests, usually in a multiple-choice format, are very common on the Internet, although they do not necessarily produce valid assessments due to the nonprofessional construction and publication of these tests. However, much effort has been made to develop online tests and questionnaires that meet rigorous psychometric standards and make valid assessments. In addition, other assessment procedures, such as online interviewing through videoconferencing, have been implemented. Counselors who operate on the Internet may use professionally established online diagnostic tests that integrate with the counseling process, thereby avoiding cumbersome paper-and-pencil testing of clients with whom they interact online. Several methods have been developed to ensure test quality and to verify clients' identities if necessary. It should be kept in mind, however, that whereas online tests produce test-based assessment, counselors miss the opportunity to diagnose clients in relation to their test-performing behaviors (e.g., nervousness).

These 10 principles characterize Internet counseling and make it unique and different from any previously practiced psychological interventions. These principles represent a significant revolution in the traditional conception of psychotherapy. Perhaps this is the reason why many have voiced concern about its implementation and growing acceptance by practitioners.

4. RESEARCH ON INTERNET COUNSELING

As a relatively new area of conduct, Internet counseling has been subjected to empirical study in only a limited scope. Yet published research reports, or those presented at public conferences, can shed some light on intriguing questions regarding therapeutic process and outcome as well as on other important subjects (e.g., training of counselors). It is important to note that the methods used for research on Internet behavior and Internet-related interventions are somewhat different from those used for research on offline counseling. In Internet counseling, individuals often are anonymous and unidentifiable, people's behaviors often parallel their written scripts (and not their observed behaviors), technical difficulties might interfere with the durability of data, imposters and fakers among participants are possible, dropout rates are higher than usual, and so forth. Research has been conducted on various questions that are important for Internet counseling, including some issues relating to the process and outcome of a variety of online counseling interventions, the impact of online support groups, and the quality of measurement of online testing. In addition to research, quite a few case studies have been presented in publications and at conferences.

An overview of the empirical research on Internet counseling provides support of an emerging approach. Process-related research has found that clients of online therapy actually build therapeutic relationships, communicate and express themselves effectively, experience therapeutic alliances, and find relationships with therapists to be satisfactory. Outcome research in numerous areas of distress, such as depression, several types of anxiety, recurrent headaches, eating disorders, pathological grief, panic disorder, and general concerns, often provides positive results, usually in comparing effects with those of parallel offline treatment. Most of this

research has found that Internet counseling leads to results that are similar to, and sometimes even better than, those of traditional counseling. Feedback from Internet counseling users frequently highlights their appreciation of flexibility of use in terms of place and time, the value of experiencing sensitive (and sometimes threatening) relationships and avoiding physical proximity and eye contact, the importance of written records of conversations, and the reassuring feeling due to the options of anonymity and easy termination.

More research has been published on online support groups than on online psychotherapy. Studies have reported that online group dynamics are generally similar to those of offline groups in terms of therapeutic factors and group processes, although the weight and impact of these factors may vary. Outcome research on online support groups has consistently found a great impact on participants in terms of emotional relief and symptom reduction. Users of online support groups repeatedly praise this activity and highlight the ability to find a place where other people go through difficulties similar to their own and to share feelings and thoughts with them, receive and provide advice, and experience nonjudgmental acceptance.

Research on Internet-based assessment, especially online testing, has generally been very supportive. In most published research, it has been found that when online diagnostic tests are professionally constructed and managed, they offer psychometric qualities similar to those of their offline equivalents. Moreover, most test takers find this activity to be convenient and satisfactory. However, research does show that test parameters may change significantly from offline to online versions. Thus, a new standardization is suggested, as is caution in interpreting raw scores derived from the Internet.

In contrast to those findings, research on self-help Web sites has consistently found the validity of information provided to users to be questionable. Problems have been found regarding wrong or even misleading information, superficial and nonprofessional information, obsolete information and the slow pace of updates, and biased and selective information (in many cases reflecting Web site owners' agendas). Not only may these problems result in not helping users' behavioral and/or emotional states, they might even cause more harm in extreme situations.

Very little research has been published on software-managed Internet counseling. Although experimental data from laboratory studies seem to be promising, there is still much to be investigated in terms of actual effectiveness and the impact on real online clients.

5. ETHICAL, PROFESSIONAL, AND LEGAL CONSIDERATIONS

The special challenges introduced by online provision of psychological applications, especially treatment interventions, have generated debate, especially on ethical issues. It seems that two significant differences from the traditional counseling encounter—the invisibility of therapists and patients, on the one hand, and the communication from a distance, on the other—have triggered much criticism, at times even leading to a point of total debarment. This outcome comes on top of the reliance on new and developing technologies and unstable communication technology, many people's uncertainty and anxiety over technology in general, the lack of computer knowledge and skills, the discriminating digital gap, and the fluidity of professional boundaries, to mention just a few factors. All of these have created the need for thorough ethical and professional considerations and subsequent remedies. Several professional organizations (e.g., American Psychological Association, Clinical Social Work Federation) have published statements emphasizing the concern and need for special caution. Other professional organizations (e.g., American Board of Certified Counselors, American Counseling Association, International Society for Mental Health On-line) have endorsed special ethical codes regarding the provision of Internet counseling.

The ethical standards according to which professionals are advised to practice Internet counseling correspond to, and coincide with, general ethical guidelines for counseling practice. These standards also complement their special online environment and connection with clients. Several major principles form the center of the common ethical codes. First, because online communication is essentially different from regular face-to-face communication, clients ought to be oriented about online limitations in communicating (e.g., lack of nonverbal cues, technological failures) and how to deal with them. Second, counselors must reveal some of their personal details (e.g., full names) and share their professional qualifications with clients. Third, counselors are expected to explain to clients the potential benefits as well as the potential risks of Internet counseling. Fourth, counselors are advised to create and come to an agreement with clients on rules according to which counseling will take place (e.g., frequency of and reaction time to e-mails). Fifth, counselors ought to practice special safeguards in protecting the secrecy of clients' records and correspondence (e.g., use encrypted e-mail, chat through a secure site,

protect computer records by means of a password). Sixth, counselors should discuss emergency situations with clients, including alternative channels of communication and local backup service.

There are quite a few other ethics-related, professional, and legal issues and considerations of which counselors must be aware. The easy access to and reception of Internet counseling by minors necessitates special attention to clients' ages and acquiring guardians' legal consent (where the law requires it) in cases where minors are referred to counseling. Obviously, this is not an easy task to deal with given that many young referrals want to maintain secrecy—often from their parents—and preserve anonymity. The ability to provide counseling services to clients from other states, provinces, and countries introduces complicated legal problems that reflect issues of jurisdiction, legal accountability, and the like, leading to possible lawsuits. In addition, the ability and ease of being referred to counselors from other cultures—a very common phenomenon—necessitates a special awareness of language limitations and intercultural differences that might significantly intervene in the counseling process and outcome. Standard malpractice insurance might not cover online professional practice; hence, counselors must take special cautionary steps.

Another issue deals with providing counseling to anonymous clients. As indicated previously, this factor has a significant impact on the special value of Internet counseling; however, it also raises concerns of a possible legal (e.g., parental consent for minors, duty to report of crime) and professional (e.g., emergency action at a time of potential suicide or violence) nature. It should be noted here that special attention has been given to suicidal referrals and to people who are in emergency situations. Perhaps related to the harboring of traditional conceptions and habits—assuming that physical proximity and face-to-face visibility have special value in helping people and in saving lives in extreme situations—is the fact that many Internet counselors prefer not to accept clients in such situations for therapy and instead prefer to refer them to face-to-face services or to telephone hotlines. It should be noted, however, that the support provided through online channels to suicidal people and to people who were extremely disturbed emotionally has proven to be as successful as that provided through offline traditional services. Owing to anonymity, ease of termination, and the lack of eye contact, it seems that people in these situations may be drawn to online counseling more easily than to offline counseling, thereby giving them a better chance of being helped.

To deal with the identity and authenticity of online counselors, a system of credentials checking has been suggested. Counselors who desire this service would supply full details and documents relating to their identities and qualifications (i.e., professional training and experience). These would be checked and verified with the authorizing bodies and institutions. When confirmed, counselors would receive special icons to be posted on their Web sites. Concomitantly, the system would educate potential users to apply only to counselors whose sites carry the permit. Although this idea might solve a major problem that currently characterizes Internet counseling, its application is still in the developmental stage.

Payment for Internet counseling is another professional issue. Although it was originally thought that fees for online counseling would be lower than those for face-to-face counseling, the fees (where counseling is not provided free of charge) are actually similar to those for in-person counseling. It seems that the special expenses for technology and continuous upgrades, the cost of required security measures, and the extensive time needed for counselors to type (as opposed to talk) made the preliminary expectations concerning reduced fees premature.

6. LIMITATIONS AND PROBLEMS

Internet counseling is not flawless or simple to use and is vulnerable to quite a few problems. The heavy dependence on technology, including electricity, the numerous components of personal computers, and complicated Internet communication, makes the whole system vulnerable and fragile. Working with sensitive clients in shaky situations under these circumstances imparts an uncertainty, or even risk, to Internet counseling. Internet counseling also requires that both counselors and clients be skilled in the use of computers and the Internet because a lack of skills or limited skills directly affects the counseling process. Unskilled counselors might, for example, lose contact with clients in the midst of highly sensitive sessions without preparing the clients for such incidents, where the consequences might be detrimental. Moreover, this type of counseling, suits those who find Internet communication to be easy and convenient, who prefer reading and writing over listening and speaking, and who do not find computer technology and Internet surfing to be intimidating. Still another problem has to do with the "digital divide," that is, discrimination

against "economically challenged" individuals. Although free Internet services are available in some public places, this certainly cannot be considered to be a solution to a problem that prevents many from receiving counseling.

The existence of sophisticated hackers and computer-savvy individuals who might infiltrate private computer zones and copy, alter, or eavesdrop on confidential materials represents another severe problem. Although developing legislation worldwide and changing the policy and priorities of the police regarding hackers might seem to lessen this problem in the future, and although offices and/or file cabinets certainly are not impenetrable themselves, Internet counseling is still considered to be less secure than face-to-face counseling. However, the growing use of encrypted online correspondence reduces the severity of this problem. Finally, a major problem has to do with untrained professionals who practice online. The practice of Internet counseling requires special incremental training, on top of standard counseling education, that focuses on the specific knowledge needed to practice counseling online. Professionals who offer Internet counseling without appropriate training might not provide optimal counseling to clients, at best, and might cause harm to them, at worst.

7. SUMMARY

Internet counseling harnesses advanced computer and communication technology so that professionals can provide consultation and therapeutic services to people who can and want to use the Internet for these purposes. Despite the criticisms voiced by individuals and organizations alike regarding problems and disadvantages of the new approach compared with traditional face-to-face counseling, this movement has been growing in size and scope. With proper awareness of the risks and limitations, appropriate professional training, strict and well-communicated ethical codes, public education, and perhaps legal actions, Internet counseling might prove not only to be useful and efficient but also to exceed traditional counseling in many ways. The possibility of reaching remote populations that do not receive counseling services regularly and of providing convenient access to disadvantaged, shy, and disabled people as well as people who want to avoid exposure, the importance of automatically stored records for reflection and further psychological analysis, and the great opportunities for education, supervision, and research all make Internet counseling unique and useful. In the current trend, it seems that professional associations are cautiously and gradually endorsing Internet counseling as a legitimate, but not problem-free, method of providing counseling.

See Also the Following Articles

Cyberpsychology

Further Reading

Barak, A. (1999). Psychological applications on the Internet: A discipline on the threshold of a new millennium. *Applied and Preventive Psychology*, *8*, 231–246.

Barak, A., & English, N. (2002). Prospects and limitations of psychological testing on the Internet. *Journal of Technology in Human Services*, *19*(2/3), 65–89.

Bloom, J. W., & Walz, R. G. (Eds.). (2000). *Cybercounseling and cyberlearning: Strategies and resources for the new millennium*. Alexandria, VA: American Counseling Association.

Boer, P. M. (2001). *Career counseling over the Internet*. Mahwah, NJ: Lawrence Erlbaum.

Fenichel, M., Suler, J., Barak, A., Zelvin, E., Jones, G., Munro, K., Meunier, V., & Walker-Schmucker, W. (2002). Myths and realities of online clinical work. *CyberPsychology & Behavior*, *5*, 481–497.

Fink, J. (Ed.). (1999). *How to use computers and cyberspace in the clinical practice of psychotherapy*. Northvale, NJ: Jason Aronson.

Gackenbach, J. (Ed.). (1998). *Psychology and the Internet*. San Diego: Academic Press.

Goss, S., & Anthony, K. (2003) (Eds.). *Technology in counselling and psychotherapy: A practitioner's guide*. London: Palgrave Macmillan.

Hsiung, R. C. (Ed.). (2002). *E-therapy: Case studies, guiding principles, and the clinical potential of the Internet*. New York: Norton.

Joinson, A. (2002). *Understanding the psychology of Internet behaviour*. Basingstoke, UK: Palgrave Macmillan.

Kraus, R., Zack, J., & Stricker, G. (Eds.). (2003). *Online counseling: A handbook for mental health professionals*. San Diego: Academic Press.

Ritterband, L. M., Gonder-Frederick, L. A., Cox, D. J., Clifton, A. D., West, R. W., & Borowitz, S. M. (2003). Internet interventions: In review, in use, and into the future. *Professional Psychology: Research and Practice*, *34*, 527–534.

Riva, G., & Galimberti, C. (Eds.). (2001). *Towards cyberpsychology: Mind, cognition, and society in the Internet age*. Amsterdam: IOS Press.

Rochlen, A. B., Zack, J. S., & Speyer, C. (2004). Online therapy: Review of relevant definitions, debates, and

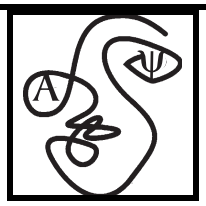
current empirical support. *Journal of Clinical Psychology*, *60*, 269–283.

Stofle, G. S. (2001). *Choosing an online therapist: A step-by-step guide to finding professional help on the Web*. Harrisburg, PA: White Hat Communications.

Suler, J. (1996–2004). The psychology of cyberspace. Retrieved on April 20, 2004, from http://www.rider.edu/~suler/psycyber/psycyber.html.

Suler, J., Barak, A., Chechele, P., Fenichel, M., Hsiung, R., Maguire, J., Meunier, V., Stofle, G., Tucker-Ladd, C., Vardell, M., & Walker-Schmucker, W. (2001). Assessing a person's suitability for online therapy. *CyberPsychology & Behavior*, *4*, 675–679. (see also correction, *CyberPsychology & Behavior, 5, 93)*.

Wallace, P. (1999). *The psychology of the Internet*. New York: Cambridge University Press.



Interpersonal Attraction

Carlos Yela

Complutense University, Madrid, Spain

1. Introduction
2. What Is IA and What Generates It?
3. Causes and Explanations for IA: Theories of Social Psychology and Levels of Analysis
4. Applications of the Psychology of IA
Further Reading

GLOSSARY

arousal Physiological activation level of a person at a given moment, which varies according to many situational and psychological variables.

complimentary of needs An interpersonal state in which one's needs tend to be satisfied by another's behavior, and one's behavior tends to satisfy the other's needs.

personally desirable characteristics Characteristics of a subject that are perceived as positive by another, regardless of whether they are socially desirable or not.

pheromones Biochemical hormone-like substances that seem to be involved in eliciting sexual attraction and perhaps also interpersonal attraction (and a promising and intriguing future area of research).

propinquity Physical proximity between two (or more) persons, which increases the possibility of interaction and therefore familiarity, with a tendency to increase interpersonal attraction.

releasers In sociobiological language, stimuli that tend to elicit an automatic, unconditioned attraction response (in contrast to conditioned stimuli, which raise attraction because of their association with a releaser or unconditioned stimulus).

socialization A lifelong process through which every person learns cultural, social, and group norms, both explicit and implicit, through the influence of parents, siblings, peers, teachers, the media, and social organizations.

socially desirable characteristics Characteristics of a subject that are perceived as positive for the majority of a society.

Interpersonal attraction (IA) plays a major role in human lives, improving both personal well-being and interpersonal relationships, or on the contrary, resulting in adaptive, interpersonal, and psychological problems if one fails to attract others or to find others attractive. The main factors that provoke IA are socially and personally desirable characteristics (including physical attractiveness), similarity, reciprocity of attraction, propinquity, familiarity, high arousal level, complimentary of needs and, presumably, pheromones. People feel attracted to both unconditioned and conditioned stimuli, thus not only are biological processes relevant, but also sociocultural, interpersonal and psychological factors all play a role. Sociopsychological knowledge can be applied to try to improve IA so as to construct better interpersonal relationships and improve quality of life. This can be done in a variety of fields and contexts, such as within organizations, sports teams, school, the family, the couple, therapy, politics, and international and intercultural relationships.

©2004 Elsevier Inc.
All rights reserved.

1. INTRODUCTION

All human beings are heavily influenced by interpersonal attraction (IA) in their everyday lives, and this has been the case throughout the history of humankind. In this threatening and problematic world, IA may be one of the keys to harmonious relationships, and its absence leaves the door open to negative feelings and negative interpersonal behavior patterns. Clearly, it would be blithe to suggest that social problems can be solved through increased IA; rather, attention should be focused on the major role IA can play in improving personal well-being and interpersonal and social relationships.

Although it contrasts in important ways, IA is firmly linked to other basic feelings and behavior patterns essential for personal well-being and everyday social interaction, such as the need for affiliation (and group processes), fondness, friendship, and love itself. Failure to attract others or to find others attractive often results in adaptive, interpersonal, and psychological problems.

2. WHAT IS IA AND WHAT GENERATES IT?

IA is usually described as a positive attitude toward someone, or an attitude or evaluation of other people in terms of how much one likes or dislikes them. Being an attitude, IA has an affective, a cognitive, and a motivational behavior component, and is mainly learned via the socialization process and in everyday interaction. Like all other attitudes, IA influences (and is influenced by) people's behavior, cognition, and feelings.

In summary, IA is based on the need for affiliation of he or she who is attracted, and its origins can be found in a feature or features of the attractor as they are perceived by the attracted. This article does not provide an exhaustive list of different factors and processes that psychological research has shown to generate IA, nor an attempt to identify those that are strictly necessary, but rather a general overview of such complex phenomena is presented, based on the main findings of empirical psychological research, whose general conclusions, despite exceptions, lead to a better overall understanding of IA. The characteristics that social psychological research has shown to generate IA, then, are the following, as shown in Table I.

TABLE I
Main Factors That Generate IA from A to B

- Socially desirable characteristics of B (perceived by A)
  - Physical attractiveness
  - Social skills
  - Intelligence
  - Competence
  - Affection
  - Kindness
  - Sense of humor
  - Generosity
- Personal characteristics of B (perceived by A) specially valuable for A
- Similarity between A and B (perceived by A)
  - Values
  - Attitudes
  - Interests
  - Customs
  - Opinions
  - Likes (and dislikes)
- Reciprocity of attraction between A and B (perceived by A)
- Propinquity between A and B
- Familiarity between A and B
- Complementarity of needs between A and B (e.g., talkative and a good listener, protective and vulnerable, etc.)
- High arousal level of A
- Pheromones of B

1. Socially desirable characteristics of the subject of attraction, such as social skills, intelligence, competence, affection, kindness, a sense of humor, and generosity. It is important not to underestimate the influence of physical attractiveness, which has been shown to have stronger links with IA and with other positive attributions than most would like to acknowledge. Whether this is due to the halo effect, to the learned association of "what is beautiful is good," or to both, the fact of the matter is that although it may seem unjust, there are clear indications of this effect.

2. Personally desirable characteristics of the subject of attraction: personally desirable for the attracted, meaning that some features not universally considered as positive are specially relevant for the attracted. Examples of such traits include self-sufficiency, modesty, seriousness, being easygoing, eroticism, femininity, masculinity, leadership, sociocultural status or prestige, knowledge, or even a particular hobby or interest, a reminder that "beauty is in the eyes of the beholder." This is one of the more interesting questions regarding IA: what leads a person to feel attracted toward one kind of person and not another? Where does attraction to specific traits come from? In this section, this question is addressed, calling on the biological, cultural, social, interpersonal,

and finally personal (psychological) forces that exert their influence on the individual's feelings of attraction toward some particular trait or person. In simple terms, the socialization process, life experiences, history of reinforcements, and everyday interaction that contribute to ones personality all offer some of the most coherent answers to these often-asked questions.

3. Similarity between attractor and attracted, not only in core values, but also in many other areas: attitudes, interests, customs, opinions, likes and dislikes, etc. This is one of the most firmly established and empirically supported findings in social psychology, although it is not always accepted by the general public. And whereas similarity generates attraction, dissimilarity tends to generate rejection.

4. Reciprocity of attraction: it is also well known that being liked by another is a good predictor for being attracted toward him or her, because people certainly like to be liked (just as being rejected is a good predictor for rejecting someone). The strength of this effect is such that the mere perception of reciprocity is sufficient to trigger IA, regardless of whether this perceived reciprocity is real.

There are some other factors that have been related to IA, such as propinquity (although the curious phenomenon of IA across the Internet is becoming increasingly pertinent); familiarity, in accordance with the "mere exposure" effect; a high arousal level, such as the one evoked by a new, risky, uncertain, or ambiguous situation, or a situation in which our senses are overloaded; or complementarity of needs, something that leads people to the often mistaken conclusion that "opposites attract." Alongside these other factors, some authors have outlined the role of biochemical substances named pheromones, which seem to be related to sexual attraction and probably also to IA. Thus, popular references to the "chemistry" of a relationship may have some real foundations. And there is one last factor that may be just as crucial to IA as it is in the rest of human lives: mere chance (though perhaps somewhat difficult for some scientists to accept). To be in the same place at the same time, with the same aims and attitudes, and to be introduced, or simply informed of the other's existence, clearly introduces a note of randomness to the incidence of attraction. IA, therefore, is one of the bases for friendship and love, but these other interpersonal relationships imply other factors outside the scope of this section.

As with any interpersonal process, IA is not a static but a dynamic phenomenon. It begins at a certain moment, develops, and can be maintained (through intermittent reinforcement, by establishing a good balance between costs and rewards, or through cognitive processes like attribution or dissonance, for example) or diminished (through the learning laws of habituation and satiation, or a bad costs–rewards balance). It is just as important to focus, then, on factors that tend to maintain or diminish IA as well as those that generate it.

3. CAUSES AND EXPLANATIONS FOR IA: THEORIES OF SOCIAL PSYCHOLOGY AND LEVELS OF ANALYSIS

There is already an overwhelming body of literature on IA, ranging from classic studies to the most recent work. In the last few decades, there has been a constant stream of research in the field, and hundreds of papers have been published. Some research is serious and systematic; some, quite frankly, is not. With reference to the former, many theories have been offered regarding IA, to explain why humans feel particularly attracted to some individuals. Some of those theories propose explanations in terms of reinforcement, others focus on cognitive processes, others still on the concepts of social exchange. Taking the findings as a whole, it could be said that an individual will probably feel IA for those who match one or more of the following conditions: those who reinforce his or her behavior, those who produce reinforcement expectations in the individual, those who show similar attitudes to him or her, those who satisfy his or her needs, those who decrease everyday worries and negative feelings, those who provide more benefits than costs to the individual, those who provide him or her approximately the same benefits as costs, those who provide just the type of resources the individual desires, those who the individual chooses of his or her own "free will," those who are attracted by the individual, those who the individual feels he or she should not feel IA for, those whose behavior is reinforced by him or her, those who reinforce the individual "freely," and/or those who give him or her positive information.

Each one of these different conditions is upheld by one or more of the many different theories of social psychology that have been applied to explain IA. Many of them are classic, well-known theories, such as theories of reinforcement (and its application to IA, such as Byrne's famous attraction paradigm), of balance, of cognitive dissonance, of social comparison, of reactance, of self-perception, of social exchange, of

TABLE II
Factors from Different Levels of Analysis That Influence IA from A to B

Level of analysis	*Factors influencing IA*
Species	Biological factors, pheromones, releasers
Culture	Esthetic criteria, cultural trends, culturally valued traits
Society	Social norms of IA, social influences, social pressures, socialization agents, and organizations
Interpersonal	Everyday interaction, group processes (conformity, modeling, etc.), propinquity, familiarity, reciprocity, complementarity
Individual	Personality (specific characteristics, values, attitudes, attraction to certain specific characteristics), physical attractiveness, needs, arousal level (generated by interaction with the situation)

gain–loss, of interdependence, and of equity. However, the process of IA, as with many other interpersonal processes, is deeply influenced by a wide range of factors, which can be understood more easily if some different levels of analysis are considered, as shown in Table II.

3.1. Species

As a member of the human species, we are often physically attracted to physical stimuli called "releasers," which tend to elicit an automatic response of attraction. Thus, sociobiological researchers have found that as a general rule, despite usually quite significant cross-cultural differences, men tend to feel more attracted by women with neotonous (child-like) facial characteristics, large and firm breasts, narrow waists, fine eyebrows, moderately wide hips, a slim and curvaceous figure (commonly reinforced by high-heeled shoes in some cultures), long and slender legs (exaggerated by the use of short skirts), a slightly protruding backside, and a lack of body hair. According to the same kind of studies, women, meanwhile, tend to feel attracted by a strong jaw, rough skin, bushy eyebrows, a firm and compact backside, a robust frame, tall stature, large eyes, and a flattish stomach. Furthermore, the role of pheromones, mentioned previously, is another biological factor related to IA. Of course, IA is not the same as physical attraction, but many social psychology researchers have demonstrated a very strong link between the two phenomena, such that people often feel IA for those who seem physically attractive to them, just as they tend to perceive those whom they feel IA for as more physically attractive than they supposedly are. Moreover, as was previously mentioned, neither IA nor physical attraction are evoked only by unconditional (automatic, instinctive) stimuli, but rather by conditioned (learned) ones, which each individual has learned to associate with that response throughout socialization processes. This is exemplified, for example, by changes in the concept of "beauty" through the ages.

3.2. Culture

IA is also strongly influenced by aesthetic criteria and cultural trends and values, which vary from one culture to another, which is to say that they also vary from one historical moment to another. In fact, one of the main factors that often precedes IA, the presence of socially desirable characteristics in the attractor, is something that depends mainly on the norms of each culture. One culture (at a given moment in time) may value the capacity to defend oneself in the face of adversity, while another may value the capacity to ask for help and community support, to give one simple example. Cross-cultural social psychology has shown us, for example, the variability between individualist versus collectivist cultures in terms of the norms and rules that prevails in interpersonal relationships.

3.3. Society

Social norms, continuously constructed and reconstructed in everyday interactions and transmitted by the socialization agents (family, school, social and work organizations, church, mass media, etc.), play an important role in the process of IA. Social factors influence what is socially desirable, as well as the configuration of an individual's personality, which in turn influences what each person considers personally desirable in others. Thus IA, as well as physical attraction and love, is not simply a psychological process, as it is commonly perceived by laypeople, but more precisely is a sociopsychological (in fact, biosociopsychological) process.

3.4. Interpersonal Interactions

Social pressures and social influence are usually exerted through interpersonal relationships, especially within the groups in which people interact every day. Group attitudes and group norms regarding other groups and other people (according to their ages, race, gender, attitudes, etc.) have a major influence on the factors that are likely to elicit and to maintain IA. Conformity, polarization, or even group thinking may all have a role to play. Readers should be aware that the behavior (actions, feelings, and cognitions) of the people with whom a person interacts every day has an important effect on that person's attitudes and behavior, including the IA he or she feels toward other people, and including the decisive role of reciprocity of attraction. Others' opinions and the pressure they exert are often quite relevant to people's feelings of IA toward someone (whether they lead to conform to their standards or to reject them by means of a "reactance process").

3.5. Individual

Finally, of course, the individual is also important. For some authors this is the main factor in any case; for others, the individual is less relevant than people would like to admit, being strongly shackled by biological and social forces. People's behavior in general, and more specifically feelings of attraction toward others, are constrained by biological, cultural, social, and interpersonal processes and factors but in truth are "acted out" or felt by a specific individual with specific psychological characteristics. Psychological studies have revealed that some features are more relevant to IA, such as physical attractiveness, social skills, or personality variables that happen to be socially recognized as valuable. In addition to these, and as mentioned previously, each person will feel attracted to certain specific features that are especially valued by the individual in question. Furthermore, one's own psychological characteristics are of considerable importance, because their similarity to the psychological characteristics of the other is one of the main factors (if not the main one) responsible for IA. In addition to these, complementarity of needs (between attractor and attracted) and personal need for affiliation are two further relevant variables affecting the IA process. Clearly, there are too many factors here to ignore, underlining the crucial role that the individual characteristics of each person play in establishing his or her feelings of IA toward one person rather than another. Nevertheless, psychological researchers are still far from establishing any kind of simple (and probably simplistic) taxonomy of types of personalities associated with tendencies to feel IA for other specific types of personalities.

4. APPLICATIONS OF THE PSYCHOLOGY OF IA

Psychological knowledge of IA not only leads to a better understanding of the human being, but also offers opportunities to apply that knowledge, and thus solve some personal and interpersonal problems, increasing subjective well-being and quality of life.

As was pointed out earlier, the absence of feelings of IA to others and/or not being the object of attraction for others will probably result in severe psychological difficulties. So, psychological knowledge about what generates IA can be used to assess and train people in these characteristics, at least to some extent. Using sociometry, social-skills training, role-playing, and other systematic and well-established psychological techniques, it is possible to assess IA and to teach people ways to make themselves more attractive, not only physically but also personally. Sociopsychological knowledge of IA can be used to give guidelines for attracting others and initiating and establishing relationships, with a reasonable degree of success, by taking into account factors such as propinquity, context, physical appearance, verbal and nonverbal communication (conversational and listening techniques, attractive body language etc.), overcoming fear of rejection, self-presentation strategies, socially desirable characteristics, accurate rhythm of self-disclosure, and so on. More constructively, psychologists can design programs conceived to increase tolerance of differences, so as not to diminish the attraction felt for someone other than oneself.

However, there is far more basic research than there are application programs for IA, and the majority of these applications are centered on the couple as the fundamental context of application. Furthermore, it is physical attraction, much more than IA, that has been applied in various fields. And, in addition, normal "applications" of IA are frequently no more than research studies concerning IA applied to a specific area (e.g., the study of the role of IA on the personnel selection process, sports team performance, jury decisions, or consumer behavior), much more than the true application of IA knowledge on solving personal and interpersonal problems and increasing quality of life.

Nonetheless, it is possible to find examples of studies concerning IA applied to certain contexts, in which those studies (and future ones) have at least a potential field of application (see Table III).

There are some well-known sociopsychological (and unconscious) processes that have been verified in some applied contexts, such as the halo effect (the influence of first impression on psychological inferences about a person), the physical attractiveness bias (the attribution of positive characteristics to attractive people and not-so-positive ones to unattractive people), and the relevance of IA on persuasion processes (IA as one of the main factors of communication agents related to success in persuasive communication). An IA halo effect has been detected in teacher expectations about his or her students (the Pygmalion effect), in falling in love, and even in contexts such as personnel selection processes and jury deliberations. A physical attractiveness bias (in some ways related to the halo effect, but not exactly the same) has been noted almost everywhere, from work organizations to everyday interactions. The impact of IA on persuasive communication is especially relevant in areas in which the capacity for influence should rely on sources other than IA, e.g., advertising, politics, or therapy.

TABLE III
Real and Potential Contexts of Application of Social Psychology of IA

Context of application	*Type of relationships*
Work organizations	Horizontal relationships, vertical relationships, personnel selection process, promotion process
Education	Teacher–students, among students
Family	Parents, parents–children, siblings
Intimate relationships	Intimate friendships, loving relationships
Therapy	Therapist–client
Sports	Manager–players, among members of a team
Courtroom	Among all figures: jury, judge, defendant, lawyer, etc.
Politics	Politicians–citizens behavior (e.g., voting)
Advertising	Salesmen and models–consumer behavior (e.g., buying)
Health	Social models—healthy/unhealthy behavior (e.g., smoking)
Everyday life	Interracial and intercultural relationships

One promising and highly relevant context of application is IA in work organizations. There is some evidence to support the influence of IA in both vertical and horizontal interactions, including the personnel selection process, and therefore in productivity and the quality of one's working life. For example, in general, data indicate a positive relationship between IA and interview outcomes, taking into account measures of personal attraction, physical attraction, and even some physical measures such as height and weight; in addition, IA moderates interactions in recruitment interviews by gender. Along the same lines, other results indicate that physically attractive applicants in jobs requiring high (but not low) levels of social competence were recommended for hire more often, were expected to perform at a higher level, and were perceived as more representative of successful employees than were unattractive applicants. It is well established that professionals are as susceptible to this bias as are college students and laypeople.

Other fields of application include the following relationships:

- between teacher and students, in which the relevance of IA has been shown to be a determinant of first impressions (and thus, as is generally accepted, on attributions and behavior), not only among students but also on the teacher's expectations and behavior toward his or her students;
- between parents and children, a field in which IA is no doubt more important than physical attraction;
- between lovers, an area in which there are real applications, such as modeling strategies to become (or to seem) more attractive to another (and, for single people seeking a relationship, psychological knowledge of IA can be applied as a serious means to form relationships);
- between therapists and clients, which might seem intriguing, given its supposed irrelevance;
- between the members of a sports team, and between players and coaches; here, IA has an effect in terms of improving results, cohesion, and satisfaction;
- between judge, jury, and defendant, with potentially dramatic consequences—for example, the fact that, in general, people tend to look more favorably on victims and defendants who are good looking;
- between politicians and citizens, IA influencing, for example, voting intention and even voting behavior, to some extent;
- between the seller (or its advertising models) and consumer, since it is well established that the IA of the

seller (or the model who advertises a product) has a major role on the success of the persuasion process for people to buy the product (here physical attraction plays a part as well as IA in general);

• between social and TV models and common people, knowing that the IA of social models (mainly perceived via TV) influences the extent to which people imitate their behaviors. This is relevant, for example, in relation to good and poor health (smoking, anorexia, etc.), violence/nonviolence, social/unsocial behavior;

• and finally, of increasing importance nowadays, there are international, interracial, and intercultural relationships in which IA is especially difficult to raise (for all of the aforementioned reasons), but at the same time is especially important in today's global environment.

See Also the Following Articles

Emotion ■ Interpersonal Behavior and Culture ■ Interpersonal Perception ■ Work and Family, Relationship between

Further Reading

Berscheid, E., & Walster, E. (1978). *Interpersonal attraction*. Reading, MA: Addison-Wesley.

Buss, D. M. (1994). *The evolution of desire: Strategies on human mating*. New York: Basic Books.

Byrne, D. (1971). *The attraction paradigm*. New York: Academic Press.

Cook, M., & Wilson, G. (Eds.). (1979). *Love and attraction*. Oxford: Pergamon Press.

Dion, K. K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality & Social Psychology*, *24*, 285–290.

Duck, S., & Gilmour, R. (Eds.). (1981). *Personal relationships*. London: Academic Press.

Etcoff, N. (2000). *Survival of the prettiest: The science of beauty*. New York: Anchor Books.

Fletcher, G. (2002). *The new science of intimate relationships*. Malden, MA: Blackwell.

Gaines, S. O., Buriel, R., Liu, J. H., & Ríos, D. I. (1997). *Culture, ethnicity and personal relationships processes*. New York: Routledge.

Hatfield, E., and Rapson, R. L. (1996). *Love and sex: Cross-cultural perspectives*. Boston, MA: Allyn & Bacon.

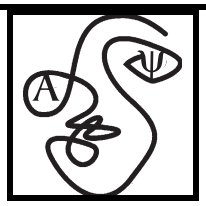
Huston, T. L. (Ed.). (1974). *Foundations on interpersonal attraction*. New York: Academic Press.

Rhodes, G., and Zebrowitz, L. A. (2001). *Facial attractiveness: Evolutionary, cognitive, and social perspectives*. Westport, CT: Ablex.

Sangrador, J. L., & Yela, C. (2000). What is beautiful is loved: Physical attractiveness in love relationships in a representative sample. *Social Behavior and Personality*, *28*, 207–218.

Wilson, G., & Nias, D. (1976). *Love's mysteries*. Glasgow: Collins Sons & Co.

Yela, C. (2002). *El amor desde la Psicología Social: ni tan libres ni tan racionales* ("Love from social psychology: Not so free nor so rational") (2nd ed.). Madrid: Pirámide.



Interpersonal Behavior and Culture

John Adamopoulos
Grand Valley State University, Allendale, Michigan, USA

1. The Interface of Culture and Interpersonal Behavior
2. Theoretical Approaches to Culture and Interpersonal Behavior
3. Thematic Approaches
4. Applications: Cross-Cultural Interaction
5. Conclusion
Further Reading

GLOSSARY

circumplex models Models of the structure of personality traits, behavior, or emotions that are characterized by a circular arrangement of their components.

cultural syndromes Patterns of behavior, cognitions, and emotions that characterize whole cultures or groups of cultures and account for psychological diversity across cultures; the most popular cultural syndromes are individualism (which emphasizes individual autonomy) and collectivism (which focuses on interdependence among group members).

distal versus proximal determinants of behavior Causes of behavior that are remote (other causes may intervene between them and the behavior) versus causes that are immediately connected to behavior.

ecocultural framework A framework for cross-cultural psychology that views human behavior both as the result of shared psychological processes and as a series of adaptations to ecological challenges.

in-group A group of people who feel strong emotional and psychological ties and who perceive that they share a common fate.

internal (dispositional) versus external (situational) attributions Explanations of one's own behavior, or of the behavior of others, that focus on the internal personal characteristics of the individual engaging in the behavior as opposed to external demands and social pressure.

interpersonal structure The psychological dimensions that describe the meaning of interpersonal relations or behavior.

resource exchange theories Theories of social behavior that view interaction as the exchange of physical and psychological resources needed for individual survival and well-being.

utilitarian models Explanatory models of human action that emphasize the purpose, function, and expected consequences of the action.

Psychological theories about the determinants of interpersonal behavior traditionally have either postulated intraindividual causes (e.g., attitudes, personality traits) or looked at the immediate social situation (e.g., group pressure, influence of the social or physical environment). However, a number of recent approaches have emphasized the role of culture in the production of behavior. In these approaches, culture has been conceptualized as a system of adaptations that transcends the individual human individual or as a subjective meaning system (or, in some instances, as both). Interpersonal behavior, then, can be viewed in part as a consequence of cultural influences or as part of a set of constructs that constitute culture. Either perspective provides considerable theoretical insight

©2004 Elsevier Inc.
All rights reserved.

into the meaning, origins, and function of interpersonal behavior and facilitates a fresh examination of established social psychological phenomena.

1. THE INTERFACE OF CULTURE AND INTERPERSONAL BEHAVIOR

There have been numerous attempts to provide a psychologically relevant definition of culture, but there is no single definition that has met with widespread agreement. One difficulty stems from disagreements between those who view culture as a purely subjective and ideational construct and those who emphasize its observable characteristics (e.g., behavioral processes). A second disagreement stems from an ontological question: Is culture a set of invented convenient categories for the classification of certain observations, or does culture have an existence of its own (i.e., independent of the individual person)?

Whatever their definition of choice, many psychologists accept the notion that culture can be viewed as a series of adaptations to environmental challenges that enhance the lives of a group of people and, therefore, are transmitted from one generation to the next. In this context, culture provides a functional basis for the analysis of interpersonal behavior. In other words, interpersonal behavior is understood as serving a purpose and as having some sort of utility in the lives of people. In a broader sense, it should be clear that any definition of culture may have profound implications for the psychological explanation of interpersonal behavior.

2. THEORETICAL APPROACHES TO CULTURE AND INTERPERSONAL BEHAVIOR

During the past two decades or so, we have witnessed a significant growth in the number of psychological theories of interpersonal behavior that are framed within a cultural perspective. Although not exhaustive, the following classification provides a sense of the direction of these efforts.

2.1. Structural Models

The pioneering work of Osgood on the measurement of connotative meaning across cultures has been an important influence on the development of structural models of social behavior. Osgood showed that three psychological dimensions—evaluation (*E*), potency (*P*), and activity (*A*)—appear to be universal (culture-common) structures involved in the understanding of the implicative meaning of things. In other words, Osgood proposed that, regardless of specific linguistic or cultural practices, humans understand the meaning of things by judging, among other attributes, how good or bad (*E*), strong or weak (*P*), and fast or slow (*A*) they appear to them.

Inspired by this work, a number of psychologists explored the structure of the interpersonal domain. The fundamental aim underlying much of this research has been the identification of a set of psychological dimensions that remain (relatively) invariant across cultures and that reveal how humans understand and communicate the meaning of interpersonal behavior. Particularly noteworthy in this effort has been Triandis's approach, which significantly influenced cross-cultural psychological research during the 1970s and later years. Triandis developed methods that could be used to study what he termed "subjective culture," that is, a group's characteristic way of understanding its social environment. This process often involved exploring belief and value systems, norms, roles, attitudes, and expectations in various cultural groups.

Central to the approach was the structural analysis of social behavior within and across cultures using a variation of Osgood's "semantic differential" technique, which Triandis called the "behavioral differential." This instrument consisted of scales gauging the perceived likelihood that people engage in different behaviors in various social contexts. Data reduction techniques were then employed to extract both culture-common and culture-specific dimensions that captured the variety of meanings of interpersonal behavior in very different cultures.

As in the case of connotative meaning, structural analyses of interpersonal behavioral intentions observed in many cultures showed an impressive convergence of findings. It appears that social behavior, regardless of what other local features may also characterize it, is understood across a wide range of cultures in terms of a few fundamental dimensions: (a) association–dissociation, (b) superordination–subordination, and (c) intimacy–formality. Specifically, the first dimension appears to gauge the extent to which an interpersonal behavior involves affiliating, approaching (as opposed to avoiding) another person, and generally giving (as opposed to denying) resources to others. The second dimension involves the notion of dominance and

control over one's social environment and gauges the extent to which a particular behavior communicates the exercise of authority and status (as opposed to submissiveness and acceptance of another's authority). Finally, the third dimension captures the extent to which a behavior involves social and physical proximity and seems to be independent of the rewarding or negative characterization of the specific behavior.

This dimensional structure, part of which has been confirmed by independent research exploring the interpersonal domain of personality, has been quite useful in summarizing the array of behaviors exhibited in daily activity and in focusing theoretical analyses of social interaction. Most significant theories in social psychology explore the behavioral domains covered by the three dimensions. For example, work on prosocial behavior or aggression clearly falls within the domain of the first dimension (association–dissociation), research on obedience and conformity falls within that of the second dimension (superordination–subordination), and the study of personal relationships falls within the domain of the third dimension (intimacy–formality).

The psychological comparison of any two or more cultures based on the three interpersonal dimensions can offer a very broad understanding of fundamental similarities and differences between or among the cultures. For example, Triandis and colleagues have shown that although research participants in the United States, India, and Japan use the three dimensions in understanding behavioral intentions in interpersonal contexts, the meaning of particular behaviors may be quite different in each of the three cultures. For example, excluding a person from one's neighborhood may imply dissociation and dominance in the United States and Japan, whereas it may be considered primarily a formal behavior in India, where the caste system plays an important role in determining the range of personal interactions.

A final question regarding this approach concerns the origins of the dimensions as psychological universals. It is generally assumed that judgments of the social environment based on the three structures offer a significant survival advantage to humans and, thus, led to their emergence over time and across various cultures. Alternative views, also based on utilitarian notions, are discussed later in this article.

2.2. Cultural Syndromes

Early work on culture and interpersonal behavior was successful in identifying a large number of cross-cultural differences (and similarities) in interpersonal behavior. For example, rates of conformity in an Asch-type situation, where group members who report an incorrect perceptual judgment induce naive research participants into also making incorrect perceptual judgments, may differ across seemingly similar cultures. However, this work was relatively unsuccessful in explaining such differences or similarities in a systematic manner. A basic problem was the relative absence of broad cultural theories appropriate for psychological research.

This situation began to change during the 1980s when a number of dimensions that could be used to describe behavioral patterns within and across cultural groups began to emerge. Hofstede's examination of work-related values in a large number of cultures provided the impetus for the development of the notion of "cultural syndromes," that is, patterns of behavior, beliefs, values, attributions, judgments, and emotions that are characteristic of whole cultures (and often groups of cultures). These patterns can be used to describe, in a systematic and predictable fashion, cross-cultural similarities and differences in interpersonal behavior. The approach has been one of the most important directions of research in the area during the past 20 years or so and has been particularly emphasized in the work of Triandis.

Whereas a number of cultural syndromes have been identified, the most successful and popular ones by far have been those of individualism and collectivism. As applied to the analysis of interpersonal behavior, individualism is a cultural pattern that is characterized by a preference for control of the social environment, personal achievement and fulfillment, self-reliance, competition, and (perhaps most important) a desire for autonomy and independence in thought, action, and emotion. Collectivism, on the other hand, is characterized by a preference for interpersonal harmony, in-group cohesiveness, and an emphasis on the interdependent nature of human relations, thoughts, and emotional experiences.

The success of these two constructs, which had been available in the social sciences for a long time before psychologists began to employ them in their work, centers entirely on their usefulness in making predictions about cultural differences in interpersonal behavior and social judgment. For example, the finding that individualism is characteristic of cultures found mainly in North America and Europe (particularly Western Europe), whereas collectivism is characteristic of Asian cultures, leads to the rather obvious prediction that social values and attitudes, as well as interpersonal behavior, will be more similar among European and

North American cultures than among European and Asian cultures.

Results of research tend to support these predictions, but not without occasional contradictory findings. In general, however, agency and independence are the building blocks of Western philosophical, value, and moral systems, whereas responsibility to community, duty, and interdependence are often the foundations of morality in collectivist cultures (e.g., India, China). Individualism–collectivism differences have also been described regarding patterns of and preferences for competition (individualism), cooperation (collectivism), and conformity (collectivism), although the particular situation (e.g., whether an interaction takes place within or outside the context of an in-group) often is critical in determining the direction of the difference. In a relevant finding in this area, collectivists have been shown to exhibit greater subordination and conformity toward in-group members, but also greater dissociation and hostility toward members of out-groups, than have individualists.

Other major differences between individualist and collectivist cultures have been identified in attributional processes. For example, collectivists are not as likely to commit what has been (mis)labeled in Western social psychology as the "fundamental attribution error," that is, the tendency of observers of behavior to rely overwhelmingly on internal and dispositional explanations for the behavior of others. Obviously, differences in attributional style can lead to differences in behavior toward others because others' behavior is perceived to be motivated by very different causes. In other words, our reactions to what others do toward us will be substantially shaped by how we tend to interpret their actions and by whether we tend to focus on personality-based or situational explanations for their behavior.

The usefulness of the constructs of individualism and collectivism has been demonstrated repeatedly during the past two decades or so. With increased reliance on these constructs, however, comes the danger of confusing description with explanation. Critics of the approach have pointed to the problem that often what is predicted as a consequence of individualism or collectivism (e.g., a behavioral difference between people of two cultures) is actually nothing more than a constitutive part of the construct itself, rendering a seemingly theoretical prediction nothing more than a tautology. With care, such problems can be avoided and these constructs can continue to provide a rich theoretical basis for the discussion of cross-cultural differences in interpersonal behavior, cognition, and the like.

2.3. Utilitarian Approaches

An implicit theme underlying much cross-cultural research over the past half-century has been the notion that much of what is observed and identified as either a difference or a similarity among cultures is functional for members of a particular group or for humans at large. This theme has either been articulated explicitly or been given a central role in several theories of interpersonal behavior.

2.3.1. Behavioral Prediction Models

Concern with the determinants of social behavior, particularly with the attitude–behavior relationship, during the 1960s and 1970s led to the development of several models that boasted fairly good behavioral prediction. The models' predictive success has been demonstrated in a variety of applied settings, from advertising to fertility-related practices, in a number of different cultures. These models, particularly those formulated by Fishbein and Triandis, have a number of common features. For example, they all treat attitudes and other intrapsychic variables as major causes of interpersonal behavior, but they also emphasize the influence of the social environment in determining behavior. Invariably, the models assume that this latter influence is filtered through internal processes. For example, the effect of others' opinions is thought to lead to the formation of personal norms of appropriate behavior. In turn, the norms are expected to affect the formation of behavioral intentions or self-instructions to engage in particular social behaviors.

These models have an explicitly utilitarian character because they consider expectations of the consequences (negative or positive) of a behavior to be among its most important determinants. The models are not always in agreement regarding the precise mechanism through which perceived behavioral consequences affect behavioral outcomes. For example, Triandis's model treats this influence as an independent process, whereas Fishbein's model embeds the influence within the attitude formation process. In either case, however, a utilitarian component of human social behavior is explicitly taken into consideration. In other words, these models rely on a combination of utility and rationality to offer explanations of interpersonal behavior.

Cultural influences are not considered explicitly in such models. However, there is a clear assumption that the set of beliefs, expectations, and personal norms that

function as proximal determinants of behavior are themselves influenced substantially by cultural inputs. Thus, culture is viewed as an important but distal behavioral determinant. For example, cultural beliefs about the role of a woman in the family, coupled with local religious or moral values about contraception and conception, may lead to the formation of personal norms that influence a woman's decisions about the size of her completed family. It would be expected, of course, that such norms would play a far more significant role in determining the behavior of collectivists, whereas personal wishes and attitudes would be more influential determinants of the behavior of individualists. Indeed, there is some empirical evidence that cultural norms about appropriate behavior tend to become internalized among collectivists, presumably having an even greater effect on their behavior.

2.3.2. Resource Exchange Theories

Another class of utilitarian models has been developed to account for systematic relationships among various dimensions of interpersonal behavior. These models typically are based on the assumption that all human social interaction involves—indeed, is designed for—the exchange of physical and psychological resources. Foa and colleagues established that any social behavior appears to involve the exchange of one or more of six classes of resources: (a) love, (b) services, (c) goods, (d) money, (e) information, and (f) status. These resource classes form a circumplex model defined by the dimensions of abstractness–concreteness and universalism–particularism. The first dimension refers to the material (concreteness) as opposed to the symbolic (abstractness) nature of a behavior, whereas the second dimension refers to the extent that the relationship of the actor to the target of the behavior plays a significant (particularism) or trivial (universalism) role in the satisfactory completion of the interaction.

Extensions of this work have led to formulations that attempt to explain the emergence of the culture-common psychological dimensions of social behavior reviewed earlier and of cultural syndromes such as individualism and collectivism. Adamopoulos has argued that the dimensions of concreteness and particularism, along with the question of whether a resource is given or denied, form fundamental constraints on human interaction and are components of a differentiation process that yields uniquely defined features of interpersonal meaning. For example, superordination is defined as the denial of abstract and particularistic resources. In this fashion, the interpersonal dimensions discussed earlier may be conceptualized as emerging from this process of constraint differentiation over long periods of time. Adamopoulos extended this idea in 1999 to account for the emergence of individualism and collectivism as cultural patterns of interpersonal behavior. In this case, an additional constraint, the orientation of a particular behavior to secure resources for the self as opposed to another person, was included in the theoretical model.

In general, these resource exchange models attempt to account for two major problems: (a) the fundamental structure of interpersonal behavior and the interrelationships among its major dimensions and (b) the emergence of interpersonal structure as a function of basic social constraints. This orientation is necessarily of a universalistic nature. In other words, these models tend to focus on themes that transcend cultural boundaries. At the same time, resource exchange models can be sensitive to cultural differences by taking into consideration the extent to which various cultures expose their individual members to different constraints. It is assumed that such differential exposure leads to behavior systems with different interpersonal emphases. For example, if ample symbolic resources are available for exchange to members of a particular culture, they are more likely to engage in behaviors that involve power and superordination than are members of cultures that provide primarily concrete resources in interpersonal exchanges. The systematic description of resource availability in various cultures is, unfortunately, a task of enormous complexity and one that is likely to take a very long time to complete.

2.3.3. Sociality

Fiske's theory of sociality proposes that there are four basic types (elemental forms) of social relationships: (a) communal sharing, where individuals within an in-group share pooled resources; (b) authority ranking, where powerful leaders control and dispense necessary resources as needed; (c) equality matching, where individuals of equal status and rank enjoy comparable resources; and (d) market pricing, where interpersonal relations are a function of market demands and expectations of profit.

Fiske has argued that these four types of relationships account for most of the variability in interpersonal exchanges across very different social situations. In the context of this theoretical model, the role of a culture is to shape and present particular forms of the

four types of relationships to its individual members in ways that are relevant to their lives. For example, authority ranking may be manifested very differently in the context of a military organization than in that of a cult or gang. At the same time, the theory facilitates connections among seemingly unrelated contexts.

The relative simplicity of this theoretical model makes it quite attractive. However, critics have pointed out that, in fact, the four types of sociality deal with only a portion of the range of interpersonal relations. Specifically, the four types can easily be classified along the dimensions of association–dissociation and superordination–subordination. This, of course, indicates that this model cannot explain other aspects of interpersonal behavior.

2.4. Ecocultural Framework

The view of culture as a series of individual and group adaptations to the environment has a long history in the social sciences. During recent years, many psychologists have used Berry's "ecocultural framework" as a means of embedding such notions about adaptation into psychological research across cultures. This framework rests on two fundamental assumptions. First, humans share a number of basic psychological processes in thought and action, although such "psychological universals" appear with nearly infinite variation across the world's cultures. Second, human behavior is adaptive in that it rises in response to ecological, cultural, and situational challenges that account for the tremendous diversity evident in human life.

As elaborated by Berry and colleagues in 2002, the ecocultural framework is a dynamic system of population- and individual-level variables with multiple sources of influence. Despite its dynamic nature, however, the emphasis in the framework is clearly on the manner in which ecological and sociopolitical systems of variables affect psychological processes and behavior at the group or individual level. This is accomplished through a series of biological and cultural adaptations, although once again the emphasis is overwhelmingly on the process through which cultural adaptations are transmitted at the group level and, thus, determine the psychological functioning and behavior of individual humans.

The framework has been primarily a useful guide to research on cognitive and perceptual processes across cultures, but it has also been used to account for certain cross-cultural differences in interpersonal behavior. In general, the advantage of the ecocultural framework is that it can facilitate the exploration of behavioral phenomena by transcending standard explanations in experimental psychology. Such explanations typically revolve around intrapsychic or, at best, interindividual processes. For example, personality traits, attitudes, and the presence and behavior of others (i.e., the immediate social situation) are the most typical determinants of behavior found in psychological theory. The ecocultural framework focuses attention on the role of physical resources and subsistence economies in determining the range of human activities. Often, such accounts can be quite insightful and add considerable depth to psychological explanation.

3. THEMATIC APPROACHES

Traditionally, the investigation of interpersonal processes and behavior has been organized topically due to the large number of theoretical perspectives available and the inability of any single theoretical model to account for the variety of the observed phenomena. A comprehensive review of cross-cultural research on interpersonal behavior is beyond the scope of this article. Instead, this section outlines four major areas of relevant research and, in each case, provides examples of how cultural variables may contribute to better understanding the behavior of interest.

3.1. Social Norms and Conformity

Norms, or rules for appropriate behavior in particular social contexts, demonstrate the influence of culture on interpersonal behavior because they reflect the opinions and actions of others. Norms can be internalized and become embedded in individual belief systems, particularly among collectivists. Thus, it is expected that conformity to norms varies significantly across cultures.

Considerable research, much of it done in the context of the ecocultural framework, has demonstrated the validity of this hypothesis. Specifically, consistent differences in conformity to group norms have been found between cultures with highly structured and rigid interpersonal relationships (high conformity) and cultures with considerable autonomy in how individuals fulfill their role obligations (low conformity).

A plausible, but partially speculative, explanation of such differences rests on fundamental ecological and structural differences between the two types of cultures. Cultures in which the ecology favors the emergence of agricultural economies must develop vertical

authority systems in which the presence of powerful and respected leaders facilitates the coordination of human activities involved in agricultural production. In turn, this process necessitates the development of socialization practices that encourage conformity to authority figures and social norms. In contrast, ecologies that favor the development of, for example, subsistence economies based on hunting (an activity that requires much less cooperation among many people) may lead to a relatively greater emphasis on individual autonomy during socialization and, thus, to less pressure to conform to group norms in social life.

3.2. Prosocial Behavior

Cross-cultural differences in prosocial or positive forms of interpersonal behavior have been demonstrated consistently. These differences range from being willingness to mail a letter for someone else, to asking for directions, to helping a disabled person cross the street. In fact, the variety of helping behaviors studied over the past 30 years or so has been limited only by the ingenuity of the investigators and implementation constraints.

Research findings in this area are not easy to interpret because the specific cultural characteristics that may account for prosocial behavior patterns are not always well understood. However, a consensus appears to be emerging that the cognitive load associated with everyday life, and the resulting demands on people's attention, may explain the relatively higher likelihood of helping in rural (as opposed to urban) and poor (as opposed to affluent) areas.

Other cultural characteristics may also account for observed rates of helping in interpersonal contexts. For example, cultures differ in their traditions regarding how one behaves toward strangers or foreigners. Thus, it is possible that in some cultures the influence of norms for helping others may be moderated by the relationship between the person requesting help and the person offering it. In some collectivist cultures, where social interaction is guided by norms of interdependence and where people often rely on each other for the acquisition of resources necessary for survival, interpersonal behavior follows particular cultural scripts. Such scripts inform people's notions about responsibility and duty, and they lead to expectations that others will be of help if the circumstances demand it.

3.3. Aggression

Aggression is generally defined as behavior intended to harm another person physically or psychologically. The study of aggression has been of great interest to social and biological scientists of diverse persuasions and has resulted in a number of universal findings. Two of the best-known findings are as follows. First, aggression and violence are truly ubiquitous phenomena but are manifested in very different ways and occur at different rates around the world. Second, with very few exceptions, boys and men tend to behave more aggressively across cultures than do girls and women, respectively.

The precise processes that explain these findings are not clearly understood. Biological mechanisms have been proposed on occasion, but the supporting evidence is typically not very strong. Most psychologists have instead relied on a variety of sociocultural processes. For example, child-rearing practices, socialization, and social learning have often been evoked as explanations for both male–female and cross-cultural differences in the incidence of aggressive and violent acts. It appears that, in many cultures, men not only learn to engage in more aggressive behavior than do women but also learn to interpret aggressive behavior differently than do women. Furthermore, as documented in recent work on "cultures of honor," the incidence of violence may be higher in some cultures because their members are socialized to protect themselves and their possessions. This results in increased tolerance for aggressive behavior because such behavior is viewed as serving a self-protective function.

The preceding processes all explain aggression in terms of normative influences. A contrasting perspective may also be formulated, although it is one that receives scant attention from psychologists. According to this view, aggression may be more likely to occur in those domains of interpersonal interaction for which there are no strong sociocultural norms of appropriate behavior. In such contexts, aggression becomes instrumental for individuals in acquiring certain classes of resources. The investigation of this intriguing proposal necessitates a fundamental shift in research orientation away from social influence processes and toward the manner in which humans attempt to establish agency within a complex and often threatening social environment.

3.4. Interpersonal Relationships

The study of relationships between people has been the focus of research primarily in European and American psychology. The impetus behind this work has been a growing concern with intimacy, personal relationships,

attraction, love and marriage, and dissolution processes in relationships. In all of these areas, culture is thought to play a significant role, but in many instances, actual empirical research is lacking. For example, we may understand the mythology that sustains (and often destroys) American and European personal relationships, but similar research in non-Western cultures is practically nonexistent.

Intimacy has been traditionally defined as self-disclosure that is largely a rewarding experience for two persons. The negative aspects of intimacy as a psychological phenomenon have generally been ignored or considered only in the context of relationship dissolution processes. An alternative view of intimacy is that it is the exchange of particularistic and concrete resources. In this context, intimacy is considered a universal phenomenon whose manifested form and functions are shaped and modified by culture. For example, collectivists tend to have few but highly intimate relationships, whereas the reverse is true for individualists. It is possible that this is because the greater availability of social norms in collectivist cultures makes intimate interactions more concrete experiences. In contrast, individualists must search for appropriate information to clarify the social situation. Thus, developing intimate relationships may require more work for individualists and may explain why individualists tend to consider a romantic component critical for many intimate relationships. The strong kinship and community ties that often bind collectivists may diminish the importance of romantic love in personal relationships.

Interactions involving self-disclosure that have traditionally dominated Western conceptualizations of intimacy are, by definition, based on the rather abstract process of searching for information to better define the situation. Such conceptualizations are clearly inadequate in light of the better-defined and predictable nature of interpersonal exchanges in collectivist cultures. Consequently, alternative views of intimacy that also take into account systems of social interaction from non-Western cultures must be developed.

4. APPLICATIONS: CROSS-CULTURAL INTERACTION

The theoretical approaches to the analysis of interpersonal behavior reviewed earlier typically provide explanations not only for similarities but also for differences in the actions in which people engage during their everyday lives. Thus, these approaches can offer considerable insight into understanding a variety of situations in which people with different cultural backgrounds interact with each other.

Psychologists have studied cross-cultural interaction extensively and from diverse perspectives that have emphasized both its positive and negative consequences. For example, Berry and colleagues have developed models of psychological acculturation and the effects of cross-cultural contact on individual psychological processes. These models specify some of the conditions under which persons who come into contact with another culture may select to replace their cultural identity with that of the new culture (assimilation), may choose some form of biculturalism (integration), or may reject either the new cultural identity or both the new and the old cultural identities.

Other analyses of cross-cultural interaction have focused on the concept of adjustment, that is, the ability to develop and maintain psychological well-being even in the context of interpersonal exchanges with considerable conflict and communication difficulties. Research has established that successful adjustment to changing conditions following cross-cultural contact improves performance on a variety of tasks and, of course, decreases the amount of stress typically associated with cross-cultural experiences.

Cross-cultural interaction often involves exchanges between people who rely on very different value systems to interpret each other's behavior. As mentioned earlier, there appear to be some fundamental differences in the manner in which individualists and collectivists make attributions about social behavior. Specifically, individualists seek the causes of behavior in people's personalities, attitudes, and belief systems (i.e., internal attributions), whereas collectivists tend to prefer attributions that involve the experience of social pressures, obligations, duty, and group membership (i.e., external attributions).

Interaction between members of these two distinct types of culture may be problematic and may lead to substantial misunderstandings. For example, an individualist is likely to attribute a perceived lie in an interpersonal exchange to some flaw in the actor's personality, whereas the (collectivist) actor may attribute the lie to the constraints of the social situation (e.g., an obligation to be polite and respectful). The need to resolve conflicts of this sort—quite common in cross-cultural contact—has led to the

development of a number of instruments and programs for cross-cultural training during recent years. One of them, the culture assimilator, has been shown to increase the chances that people from different cultures will make similar attributions in various social situations.

5. CONCLUSION

The influence of culture on psychological processes is nowhere more evident than in the study of the generation and specific manifestation of interpersonal behavior. At the same time, the task of explicating the role that culture plays in behavioral processes is enormous given the range of behaviors that humans produce and the variety of processes necessary to account for them.

A number of theoretical models are available to accomplish this task, each with its unique strengths and weaknesses. Often, these models make very different assumptions about the manner in which culture is involved in the production of interpersonal behavior. This diversity of assumptions may actually be helpful at this stage given that we still know relatively little about the culture–behavior link. What is needed is a cultural theory appropriate for psychology with considerable power for description, prediction, and explanation. Such a theoretical system does not exist at this point.

See Also the Following Articles

Aggression and Culture ■ Cross-Cultural Psychology, Overview ■ Cultural Syndromes ■ Prosocial Behavior, Development of

Further Reading

Adamopoulos, J., & Kashima, Y. (Eds.). (1999). *Social psychology and cultural context*. Thousand Oaks, CA: Sage.

Berry, J. W., Poortinga, Y. H., Segall, M. H., & Dasen, P. R. (2002). *Cross-cultural psychology: Research and applications* (2nd ed.). New York: Cambridge University Press.

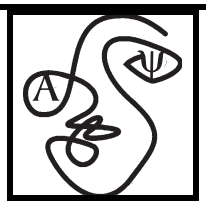
Fiske, A. P. (1990). *Structures of social life: The four elementary forms of human relations*. New York: Free Press.

Segall, M. H., Ember, C. R., & Ember, M. (1997). Aggression, crime, and warfare. In J. W. Berry, M. H. Segall, & C. Kagitcibasi (Eds.), *Handbook of cross-cultural psychology*, Vol. 3: *Social behavior and applications* (2nd ed., pp. 213–254). Boston: Allyn & Bacon.

Smith, P. B., & Bond, M. H. (1999). *Social psychology across cultures* (2nd ed.). Boston: Allyn & Bacon.

Triandis, H. C. (1994). *Culture and social behavior*. New York: McGraw–Hill.

Triandis, H. C. (1995). *Individualism and collectivism*. Boulder, CO: Westview.



Interpersonal Conflict

Panos S. Kordoutis
Aristotle University of Thessaloniki, Thessaloniki, Greece

1. Understanding and Managing Conflict in Interpersonal Relationships
2. Relationship Features and Psychological Processes Involved in Conflict
3. Overview of the Conflict Mechanism: Directives and Keys to Managing Conflict
4. Managing Conflict
5. Conclusion
Further Reading

GLOSSARY

correspondence of outcomes The degree to which the outcomes of one partner in an interdependence relationship can be attained without preventing the other partner from attaining his or her outcomes; perfectly correspondent outcomes lead to pure coordination, whereas perfectly noncorrespondent outcomes lead to pure conflict.

dispositional attributions Identifying as causes of behavior internal factors and processes such as intentions, motives, and personality traits.

interdependence or outcome interdependence A social situation involving two or more individuals who depend on one another for achieving individually valued outcomes and goals.

interference Describes Partner A's actions in obstructing Partner B's effort to bring about a personally desirable outcome; interference may be actual or perceived and may be intentional or unintentional.

interpersonal skills training Developing skills that are pertinent to constructive conflict resolution such as building problem-solving relationships with other disputants, seeking multiple solutions to conflict situations, and viewing conflict issues from an "outsider's" perspective.

maximum minimum The best possible of the worst outcomes that one is willing to accept in a negotiation setting, where agreements can be attained only by lowering both disputants' outcome levels.

mediation The intervention of an independent third party in a dispute to facilitate disputant interaction, identify sources of conflict, and promote constructive communication toward a mutually satisfying resolution.

minimum maximum The lesser among the best outcomes that one is willing to accept in a negotiation setting where disputants cannot attain their optimal outcomes.

outcome A general term employed to indicate how drives and motives shape preferences and goals worth pursuing in social situations of interdependence; outcomes may be material possessions (e.g., money, a house, a car) or abstract states (e.g., social status, security, happiness).

relative deprivation A perception that outcomes in some realm that is valued and important to self are inferior to a reasonable standard; relative deprivation is a potential source of conflict.

social comparison Using others as a comparison standard in evaluating own qualities such as efficacy, ability, judgments, and attitudes.

social norm A consensually accepted rule about behaving in social settings that has been internalized by individual society members; breeching of a social norm by an individual is sanctioned but may also cause pangs of conscience.

Interpersonal conflict is a process triggered when one of the parties in a relationship perceives that the other adversely interferes with the attainment of own

©2004 Elsevier Inc.
All rights reserved.

outcomes and goals. It may also be a consequence of perceiving own outcomes as severely inferior to those of the other in the relationship and to a commonly accepted standard.

1. UNDERSTANDING AND MANAGING CONFLICT IN INTERPERSONAL RELATIONSHIPS

Interpersonal conflict is a by-product of individuals' need to establish and maintain relationships for their intrinsic wellness and survival value. Relationships ensure that personal goals, such as a comfortable life, a rewarding professional career, security, fulfillment, and love, can be attained to a qualitative and quantitative level that either would have been impossible outside a relationship (as in the case of love) or would have required much more effort, time, and resources. Despite their motivation to uphold relationships, individuals will often come into conflict with their partners as the realization of their respective goals clash, impede one another, or become the medium of an unfavorable comparison with the other's achievements. Conflict can be a painful process, with destructive consequences for the well-being of individuals and for the potential of relationships to yield valuable outcomes for their members and society. Of course, many conflicts cause little rancor and are resolved peacefully. Intimate couples, friends, roommates, and office colleagues can often find mutually satisfying ways in which to eliminate their differences. In professional settings, managers and employees can frequently settle their disputes through simple discussions or, when matters get more complicated, through mutual demands with official labor–management negotiations. However, there are times when people will resort to overt or subdued hostility, or to verbal or physical aggression, so that they can force the resolution of differences in their favor. Hostility in all forms will also occur as a means of expressing displeasure about what is perceived as an unfair resolution of relationship differences. In either case, the ambience of personal or professional relationships is severely disturbed. Simple human interaction becomes unproductive and harmful not only for the disputants but also for those directly or indirectly related to them. For example, consequences of persistent conflict often involved in divorce are numerous and distressing for the personal development of the disputants and their children. Unfortunately, divorce rates involving intense conflict in the United States, in European Union member countries, and in other Western countries continue to be high. In professional settings, even subdued conflict between fellow employees can undermine the productivity and creativity of an entire work unit, adversely affecting the quality and quantity of products and services. Considering how threatening conflict can be to the stability and productivity of relationships, individual well-being, and the cohesiveness of society, it is not surprising that it has attracted a lot of scientific theory and research.

Social psychology, in particular, has dealt with conflict in an effort to understand it and identify ways in which to manage it. Because managing conflict requires knowledge and understanding of conflict, social psychologists have tried to develop a theory that would allow the prediction and control of conflict. The agenda of such a theory includes questions such as "What is conflict?," "What causes conflict?," "Under what circumstances does conflict occur?," "How is conflict expressed?," and "What are the social and psychological consequences of conflict?" If one could have answers to at least some of these questions, one would be able to perhaps predict, for instance, when conflict would be manifested and with what intensity. As such, one also would be able to manipulate the timing and circumstances of its manifestation so as to divert or even prevent it. Unfortunately, despite efforts, there is no true comprehensive conflict theory that can permit concise prediction and control of conflict behavior. Nevertheless, there are numerous theoretical approaches that have contributed importantly to a fair understanding of the variables involved in conflict. Thus, one might not be able to totally predict and control conflict, but one knows that its basic source is the motivational structure of the relationship in which the conflict parties are involved, namely, the fact that they are, to a certain degree, interdependent for achieving their valued outcomes. One also knows that other, more qualitative features of the relationship will influence how conflict will be expressed and that attributions about the causes of conflict to the other's internal dispositions and character will escalate conflict expression. The major psychological processes implicated by social–psychological theory in conflict production are "interference" and "relative deprivation." Interference is based on the perception that the significant other in the relationship interferes with own goals, whereas relative deprivation is based on perceiving that own outcomes are relatively inferior to those of the other and to a commonly accepted outcome

standard. By means of another intermediary process, social comparison, the first two may work to undermine self-esteem, and this in turn generates conflict to preserve, restore, and enhance self-esteem. This loose mechanism of conflict production, put together for the purposes of the current discussion by what is known about conflict so far, provides directives and keys with which to understand the development of conflict resolution and management methodologies. For example, the role of self-evaluation and self-esteem in the conflict mechanism strongly suggests that the aim of conflict resolution should not be limited to removing the "realistic" difference between disputants but rather should also ensure that the self-esteem of the conflicting parties is preserved, restored, and enhanced (if possible). The importance of knowing conflict in developing conflict management is also illustrated by the way in which conflict management skills have used the research finding that dispositional attributions escalate conflict. Conflict mediators and disputants are trained to eschew from articulating inferences about others' intentions in the relationship and personality characteristics.

The theoretical understanding of conflict not only has contributed to the development of conflict resolution methodologies but also has been conducive to diluting two lay misperceptions. First, most people commonly perceive conflict as a negative event of destructive consequences without paying sufficient attention to its more subtle positive effects. This is due mainly to the saliency of destructive consequences and to the survival value intrinsic in paying more attention to negative events. Second, people believe that disputants themselves cannot handle conflict and that conflict management skills cannot be learned. These misperceptions are perhaps even more destructive than conflict itself because they discourage people from training in conflict resolution and encourage people to avoid conflict at all costs, even in cases where conflict is called for and potentially beneficial to their personal interests and relationships. Indeed, research has indicated that people can be trained to handle conflict effectively, making the best out of it for their self-interests, self-respect, and self-expansion. Conflict can also be a creative and constructive process, helping in relationship functionality and development toward greater productivity and fruitfulness for both partners.

This article first explicates, by reference to theory, research, and real-life examples, the role of relationship features and psychological processes in yielding conflict. By pointing to how this knowledge is put into practice, the article then turns to analyzing major conflict management methodologies. It concludes by underscoring the main issues underling conflict escalation and conflict resolution.

2. RELATIONSHIP FEATURES AND PSYCHOLOGICAL PROCESSES INVOLVED IN CONFLICT

The basic relationship feature involved in generating conflict has to do with motivational structure, that is, with the fact that the partners' goals are interdependent. This so-called structural feature of outcome interdependence forces people to compare their outcomes, triggering social comparison. Thus, the next subsection discusses relationship structure along with the psychological process of social comparison. Social comparison, with its self-evaluative results (e.g., "I am better off than others," "I am worse off than others"), leads to two other psychological processes: interference and relative deprivation. These latter processes, being different in nature, are discussed in separate subsections. Nevertheless, their aim is the same, that is, to determine the extent to which the relationship partner is the culprit of negative self-evaluation. This section concludes with a discussion of how conflict experience and conflict expression are regulated by the qualitative characteristics of relationships.

2.1. Relationship Structure and Social Comparison

Interpersonal conflict springs from interdependence, an intrinsic characteristic of interpersonal relationships. Relationship partners depend on one another for achieving their wishes, goals, and preferred end states, generally called outcomes. Outcomes refer to material possessions (e.g., money, a house, a car) or to abstract states (e.g., social status and recognition, security, happiness, love). In addition to being attractive for satisfying specific needs, outcomes carry desirability and status value assigned by society. Thus, they contribute to identity definition and self-evaluation by means of social comparison. By comparing material and abstract possessions with those of others, people come to an understanding of where they stand in relation to others. In this way, relationship outcomes surreptitiously influence self-worth and establish potential for conflict within the relationship.

Relationship outcomes cannot be attained to the same qualitative and quantitative levels with the same ease and economy (in time, effort, and energy) outside a relationship. Hence, people are strongly motivated to enter relationships for optimizing their individual outcomes. That is not to say that individuals are indifferent to the common relationship outcomes or ignorant of the fact that others enter relationships motivated by the same individual outcome optimization goals. To the contrary, they are usually aware that no interdependent parties can ignore others in pursuing their goals and that all can influence one another's experiences, motives, preferences, behavior, and outcomes. As Lewin put it in his classical work on social conflict, relationship partners know that their "locomotion" toward attaining their goals may cross or be crossed by that of others. A lot of cognitive work needs to be carried out in planning the pursuit of own goals in an interdependence situation because others' plans must also be understood and predicted. Possibly, individuals will have to cooperate with others to maximize joint outcomes or compete for their own on the basis of some commonly agreed rules. Of course, rather than putting that cognitive effort into coordination and negotiation with others, direct and immediate access to desirable objectives may be demanded, and then conflict and fighting might arise. Success and failure will occur in predicting and planning, cooperating, and competing in conflict and fighting. In either case, there will be consequences for own outcomes as well as psychological repercussions for both self-worth and the continuing interaction of actors in the relationship. Others may facilitate, accelerate, and promote the attainment of own outcomes or, in contrast, may obstruct, delay, block, or simply interfere with their attainment by merely unilaterally pursuing their own goals in a relationship. The strong motivation of people to stay in a relationship and benefit from the advantages of outcome interdependence explains why they are willing to tolerate the nuisance of others' interference with personal outcomes.

2.2. Interference

It is reasonable to expect that people sharing the same social milieu may also see commonly desirable outcomes as worthwhile to pursue. Unfortunately, outcomes, whether material or abstract, may be limited; they might not be available to all, at all times, and in the same quality or quantity. Actually, some may be depleting fast as more people claim them for themselves and as more people claim more of them. Realizing this fact, social norms have been instituted in nearly all organized societies about coordinating, cooperating, and competing in claiming commonly valued outcomes. People also hold general implicit expectations, known as lay theories, about fair and accepted practices in pursuing such outcomes.

If partners sharing a relationship decide to pursue similar depleting outcomes, their ways are likely to cross. In that case, they might perceive each other as interfering with own outcomes, and conflict might erupt to remove interference and restore unhindered and direct access to desired outcomes. Conflict is most probable when each of the two partners insists on unilaterally claiming for self the maximum of the desired objective and eschews coordination, cooperation, or competition along norms and lay theories.

Money is an obvious example of a material outcome that depletes quickly as the needs to be satisfied in an intimate relationship increase. For instance, consider a married couple whose members must decide how to allocate their budget to satisfy common and individual objectives. Couple members have to reach an agreement on prioritizing objectives, and their first option likely would be to accept a norm about distribution of resources between the spouses. One such norm is that maximizing of joint outcomes must come first. Alternatively, they may devise their own rule of prioritizing, give up on some of their personal goals, and take turns in having money allocated to each member at different times. For some couples, particularly those whose members maintain a relatively rich profile of activities outside the relationship, budget allocation agreements are not easy to devise without minor or major conflict.

Another example of a more abstract outcome that depletes quickly and happens to be in great demand in interpersonal relationships is time. How much time should friends spend with each other, how often, and in what quality (e.g., leisurely time in doing things together, discussions about each other's plans and dreams, simple chatting on the phone)? How does a person decide how much time to spend with each friend and in what way? Social and cultural norms about this issue are variable and also depend on individual expectations about the quality of the specific friendship. Friend A's allocation of time to Friend B might not correspond to Friend B's willingness to allocate time for the friendship. Friend B might have expected more and higher quality time from Friend A and, thus, might interpret Friend A's decision as intentional devaluation of their friendship. In a way, although the valued commodity (i.e., time) seems

the same, it acquires different meanings and values for each friend because each appears to have different expectations about friendship. In reality, the friends' valued outcome called "friendship time" is noncorrespondent. If one friend could be satisfied in terms of his or her valued outcome (i.e., be given abundant time with friend), the other friend would have to be dissatisfied. Of course, this situation may give rise to displeasure, resentment, open conflict, or even a falling out.

Actors in an interpersonal setting may indeed come into conflict due to noncorrespondent outcomes. Correspondence describes the degree to which the outcomes of one partner in an interdependence relationship can be attained without preventing the other partner from attaining his or her outcomes. Noncorrespondence, then, refers to a situation where for one partner to attain his or her goals, the other partner must fail to do so. Perfect noncorrespondence of outcomes leads to pure conflict, whereas perfect correspondence leads to pure coordination. Noncorrespondence implies that interference and conflict may arise not only by the fact that the members of an interpersonal relationship pursue similar, albeit depleting, outcomes. They could just as well pursue dissimilar outcomes that happen to be incompatible. For instance, a married couple might not realize that when one spouse pursues a high-profile career while the other seeks a low-key family life, potential for conflict may be accumulating. One spouse's lifestyle generates outcomes that are noncorrespondent to those yielded by the other's lifestyle. Scarcity of time invested in home or family activities and an abundance of time spent with nonfamily make difficult the common pursuit of a low-key family life. Spouses must pursue their aspirations unilaterally, hoping that this will not negatively influence relationship maintenance. Alternatively, they could both change their lifestyles so as to be able to carry out more joint activities and enjoy common outcomes.

Other's interference with own goals may have cognitive and emotional ramifications because it makes other and own outcomes salient to self, thereby triggering social comparison. Self-evaluation may then ensue, typified by thoughts such as "Do I really merit this goal?" and "Am I able to achieve this target?" Social comparison may also yield perceptions of threat to self-worth projected onto the other such as "Her goal is to prove that I am ineffective" and "His aim is to humiliate me."

Either category of cognitions contributes to the multiplication of conflict manifestations and to conflict escalation. Removal of interference might not be sufficient to settle conflict because disputants are focusing on the perceived threat to their self-esteem rather than on the interference per se. Other's interference may have already severely damaged perception of self-worth and efficacy, and conflict may persist with the aim of restoring them.

2.3. Relative Deprivation

Paradoxically, noninterference may also contribute to interpersonal conflict. This situation is referred to as relative deprivation. People experience relative deprivation when they realize that own outcomes are severely inferior to a commonly accepted standard and that reaching the standard by own efforts is infeasible. In a situation of this sort, significant others (who meet the standard) with whom people have relationships may unknowingly aggravate people's own predicaments. Social comparison again works to underscore both inferiority of outcomes and self-debasement. The likelihood of these effects is high when a person is comparing self to a significant other with respect to outcomes having a central role in self-definition of both parties. If the relationship is a stable personal one and the social norms encourage it, the person with the poorer outcomes is particularly likely to expect the other to interfere so as to alleviate the former's own predicament. Of course, the significant other might not agree that he or she should assist and might not be willing to do so. Noninterference or lack of assistance may cause envy and bitterness in the person with the poorer outcomes. Moreover, it might be interpreted as an active attempt by the significant other to denigrate him or her and to contemptuously seclude him or her from "what everyone has a right to have." Conflict is bound to follow. Relative deprivation as a cause of conflict is more common between social groups than between persons given that people generally either find themselves in social settings with people whose level of outcomes is similar to their own or tend to choose significant others with similar outcomes rather than dissimilar ones.

2.4. Experiencing and Expressing Conflict: The Role of Qualitative Relationship Features

Some manifestations of conflict remain discreetly in the background until conflict has fully erupted. However, most of them are not discreet at all and need no introduction. At times, all people have

experienced behavioral manifestations such as arguments, intense disagreements, antagonism, fights, and even physical violence against other people. They have experienced the emotions involved such as apprehension and mistrust as well as being irritated, hostile, angry, bitter, resentful, sullen, and ominously silent. All people have also ruminated on thoughts regarding the causes of other people's behavior, the ways in which other people interfered with their own wishes and goals, their intentions in so doing, and their predispositions and characters. People probably have experienced conflict in their relationships with acquaintances, friends, family members, spouses, and lovers as well as in their professional or more formal relationships.

Overt behavioral manifestation of conflict and the subjective experiences involved, emotions and cognitions, may differ both quantitatively and qualitatively depending on the relationship type. Relationships have been characterized in many ways. The previously discussed feature of outcome interdependence is primarily used to characterize the motivational structure of relationships because interdependence of outcomes is intertwined with the potential of developing conflict in relationships. Thus, relationships of high interdependence, such as intimate ones, are very likely to yield conflict. The more individuals depend on one another for more of their positive outcomes, the more likely they are to interfere with one another's positive outcomes. High interdependence provides a lot of opportunities for conflict while at the same time motivating people to insist on claiming their relationship outcomes because (a) they value the relationships in themselves (i.e., what they can derive from the relationships specifically), (b) they might not have alternative options (i.e., alternative relationships offering just as good outcomes), and/or (c) they might not be willing to invest in the costly procedure of seeking and constructing new relationships.

However, conflict expression per se seems to be influenced by three other qualitative (rather than structural) dimensions of relationship features: formality, stability, and personal/task orientation. More specifically, a relationship may be formal or informal, stable or transient, and personal or task oriented. A formal relationship is governed by strong social norms providing salient guidelines about appropriate interpersonal behavior. In contrast, although an informal relationship may also embed implicit behavioral guidelines, it is less directing and leaves more room for partners themselves to define interaction rules. A transient relationship is temporary and tends to occur with irregular timing, often in different places. To the contrary, a stable relationship lasts longer, has relative time regularity, and tends to occupy identifiable spaces. A task-oriented relationship is established for the sake of accomplishing specific tasks, whereas a personal relationship is initiated mainly for its own sake and the outcomes are accrued through the relationship. Communication in the former focuses on tasks and actions, whereas communication in the latter focuses on the people and their dispositions.

An example of a relationship type that is highly interdependent for outcomes while being informal, stable, and personal is, of course, the married couple. Married couples may experience conflict over any number of issues, including budget allocation, division of household chores, family roles, failed mutual expectations, jealousy, neglect, and infidelity. In fact, it appears that the more a person knows another person, the closer the relationship is and the more likely the partners are to have had experienced a broader spectrum of all the preceding behaviors, emotions, and cognitions to different degrees. Consider the common example in married couples counseling of neglect that may give rise to particularly intense and expressively "rich" conflict. When one member of the relationship maintains a high level of rewarding life outside the relationship, allowing that area to absorb more and more of that person's interest, energy, and time, the other member may start perceiving the gap between past level of personal rewards from the relationship and present outcomes to grow. Failed expectations, deterioration of outcomes, and the inevitable social comparison with the significant other will soon trigger perceptions of interference with own relationship aspirations and a sense of relative deprivation. Feelings of frustration and self-deprecation may follow, and conflict behaviors on the part of the "underbenefited" member to restore level of outcomes and reclaim loss of self-esteem will be manifested. Because the focus in close relationships tends to be on the partners and their dispositions, attributions will be made about the "meaning" of the other's interference or lack of interference (i.e., relative deprivation). Representative thoughts could include "This is another sign that her feelings about me have changed," "He does not care about me," and "I am unable to attract her anymore." Dispositional attributions escalate conflict, and removal of interference might not be sufficient to settle it because the disputants are then focusing on perceived intentional threat to self-esteem rather than on the outcomes of their relationship. Conflict in such cases may be constructive

in the sense that a relationship member is painfully made aware of the fact that he or she is entangled in a relationship that does not accrue positive outcomes to him or her anymore. Restructuring or even leaving the relationship may be necessary to extinguish negative outcomes. Nevertheless, high interdependence, as well as the possible misperception that there are no alternative solutions, might not allow an individual to see conflict as a positive and constructive opportunity. Third-party assistance may be required in this situation.

The adage that the more people know one another, the more likely the conflict, in no way implies that conflict cannot erupt between strangers and acquaintances. More often than not, conflict will occur within a transient formal or informal relationship, that is, with people who one hardly knows or who one never actually meets, for example, the inconsiderate driver who will not allow one's car to overtake his car despite one's desperate signaling and horn honking that one is rushing to an emergency or the "high hat" assistant at the store in the posh side of the town who will answer one's naive questions with a polite yet reluctant and indignant smile. Expression of conflict in transient relationships may include a different, possibly more limited, and yet more extreme set of behaviors compared with those identified in other kinds of relationships.

The anonymity of informal transient relationships may diminish respect for social norms. The brevity and superficiality of the contact with the unknown others, the nonsaliency of social norms in such situations, and the expectation that the interactions will not be followed up all may result, by means of disregarding the human qualities of the others, in particularly acute and exceptional expressions of conflict behavior involving verbal aggression or even physical assault. The anonymous crowds at athletic events have often promulgated intense expressions of interpersonal conflict among spectating funs. Similarly, continuous movement and speed, and the impersonal indirect communication through signaling and horn honking, may precipitate antagonistic behavior between drivers on highways. Extreme behavior is not as likely in formal transient relationships, such as the encounter with the store assistant, because the environment is rich with salient norms of conduct dictating limits to appropriate "protest" behavior and sanctions for inappropriate behavior. There is less anonymity, and the likelihood that the interaction with the same person might be followed up is not negligible. As a consequence of these factors, conflict expression in formal transient relationships is usually subdued and restrained.

What about formal stable relationships such as work and professional relationships? These bear a characteristic that could breed conflict, that is, increased interdependence for outcomes. On the other hand, these relationships usually include a strong set of norms due to their formality. Because they are also stable, individuals cannot discount the future consequences of their negative behaviors. Finally, because these relationships have survival value for the individuals involved, they cannot be easily abandoned for other relationships and so members are particularly motivated to retain them.

Consider a small health care institution that hires two psychologists without a clearly delineated profile for each one's duties. In such a work setting, one is given the chance to define his or her exact job profile and to enjoy the work while gaining recognition and credit toward future promotion. Hence, the two psychologists are likely to compete in demonstrating effectiveness, diligence, and expertise—a productive state of affairs for both the institution and its patients. However, assuming that this small institution has only one higher status position for a psychologist, competition for recognition and credit may soon degrade to antagonism. In patient assessment multispecialist meetings, the two may try to undermine each other's input. They may selectively use scientific knowledge and data that counter each other's specific analyses just to raise generalized hard criticism of each other's approach. Alternatively, they may resort to side remarks and whispers to other members of the team or to personal comments about the rival colleague's personality and way of working. The result is, of course, detrimental to the productivity of the whole team and potentially threatening to patient care. The team must divert attention away from case analysis to the arguments of the two psychologists and must invest time and energy in deescalating conflict and refocusing on the task at hand.

3. OVERVIEW OF THE CONFLICT MECHANISM: DIRECTIVES AND KEYS TO MANAGING CONFLICT

Interference and relative deprivation are the two basic psychological processes involved in conflict outbreak. Both processes implicate the self, particularly self-evaluation and self-esteem, because they activate social comparison with the significant other in the relationship and with a commonly accepted outcome standard. Thus, conflict is not solely about securing for self the

valued resources of interpersonal relationships but also is about enhancing, defending, and restoring self-esteem. Social comparison is a third intermediary process that acts as the interface of the other two to the self. Interference and relative deprivation lead to social comparison with the informational input that the self requires to understand where it stands in relation to others and to evaluate own abilities, achievements, social status, and more abstract personal end states such as satisfaction, happiness, and well-being. If the result of such social comparison is negative self-evaluation, conflict behavior is likely to arise because the self implicates the other in obstructing attainment of own outcomes (i.e., interference) or in nonintervening to assist in attaining better personal outcomes (i.e., relative deprivation). The saliency of self-evaluation, brought about by social comparison, diverts attention of relationship parties away from realistic conflict issues to dispositional attributions about each other's behavior. These attributions usually implicate the other in intentional threats to own self-esteem and are highly conducive to conflict escalation. All three processes are inconceivable outside a relationship that is desirable in its own outcomes by both parties. Underlying these processes is a relationship structure of outcome interdependence. The members of the interpersonal dyad depend on one another for their valued outcomes, and it is this very dependence that motivates them to remain in the relationship despite the hurdle of interference and the increased risk of relative deprivation. Therefore, a paradox results. Conflict behavior is more probable, and often more enduring, the greater the interdependence in the relationship. How conflict will be manifested, however, will be greatly influenced by the more qualitative features of the relationship. The specific repertoire of behaviors used to manifest conflict—their aggressiveness, intensity, persistence, and other behavioral characteristics—will be regulated by the formality or informality, stability or transience, and orientation (task vs personal) of the relationship.

The preceding overview of the most important concepts that have been implicated by research and theory in generating interpersonal conflict also attempts to assemble a loose framework for a general understanding of the psychological mechanism yielding conflict behavior. The framework, depicted in Table I, connects relationship features to the psychological processes of interference and relative deprivation through social comparison. The core cohesive concept mobilizing the supposed psychological mechanism of conflict production is self-esteem, that is, the need of the self to self-enhance by self-evaluating and comparing with others.

This loose mechanism of conflict production provides basic directives in conflict resolution and several keys to conflict management methodology. First, it suggests that in conflict resolution, the ultimate aim is to protect, preserve, restore, and (possibly) enhance the self-esteem of the parties involved in conflict. The "realistic" goal of eliminating the source of interference or relative deprivation should be viewed only as a step toward self-esteem preservation or enhancement. Attempting to eliminate interference and relative deprivation without taking into consideration that this action reflects on the self-esteem of the disputant parties will only temporarily appease the conflict. In conflict management methodology, skills development is inconceivable without taking into account the role of dispositional attributions in conflict escalation. Among the most important negotiating skills is avoiding allusions to the personality and intentions of disputant parties (e.g., "name calling," "reading between the lines"). A third party intervening to assist in resolving conflict cannot possibly proceed to realistic problem-solving propositions that satisfy both parties without a full analysis of the motivational structure in the relationship. The analysis should show the different ways in which each disputant's valued outcomes could be attained to a maximum possible level of individual desirability considering that the other wants to do the same. The intervening party should identify the maximum minimum (i.e., the best possible of the worst outcomes that one is willing to accept in a negotiation setting) and the minimum maximum (i.e., the minimum among the best outcomes that one is willing to accept in a negotiation setting). Furthermore, the mediator should have a good understanding of the dimensions used by conflicting parties to self-compare with the other in the relationship. Which outcomes are significant to each party, and why are they used as referents of comparison with the other and (possibly) with a common social standard? What are the consequences of such comparisons for a party's self-esteem? The nature of interference should also be clearly understood. How do the parties interfere with each other's outcomes, and with what costs and benefits for self and other? How will changes in the interference pattern change the cost–benefit ratio for self and other? Answering such questions can provide an indication of how motivated and willing the parties will be to shift their behavior to less interference or noninterference during negotiation.

TABLE I
Relationship Features and Psychological Processes Involved in Interpersonal Conflict

	Processes		
	Intermediary	*Basic*	
Relationship features	*Social comparison*	*Interference*	*Relative deprivation*
Structural–Motivational		**Informational input**	
Outcome interdependence (high or low)	Self-evaluation becomes salient. Outcomes are compared with other's and with common standards. Result is input to "basic" processes	Perception of other as interfering with own outcomes Self and other's outcomes	Perception of self-outcomes as inferior to other's and to common standards
		Mediating cognitive–affective output	
		Self-attributions, self-devaluation (e.g., "Do I really merit this goal?")	Self-attributions and self-debasement
		Dispositional attributions about other (e.g., "His aim is to humiliate me")	Dispositional attributions about other (e.g., "She intends to deprive me of what everyone has a right to have")
Qualitative Stability–transience Formality–informality Task–personal orientation		**Behavioral output: Conflict expression** Conflict expression is regulated in terms of persistence, intensity, restraint, and behavioral repertoire by the qualitative features of the relationship.	

4. MANAGING CONFLICT

The fact that conflict is more or less natural to interpersonal relationships perhaps accounts for the common perception that disputants can do little to manage conflict themselves. However, research indicates that disputants trained in interpersonal skills may be able to contain, resolve, and make the best out of conflict, turning it into a constructive and creative process. Indeed, at the individual level, when a person has the skills to handle conflict, this otherwise destructive and painful process may work to mobilize and enhance individual abilities, resources, and ingenuity. It may further motivate self-knowledge, assert self-determination, clarify personal goals and desires, and encourage self-expansion. At the dyadic level, when each of the disputants possesses the skills to handle conflict (to a greater or lesser extent), the disputants may turn a poor interdependence relationship into a fruitful one for both parties. It appears that when individuals appreciate their outcome interdependence, conflict may provide the thrust that pulls together resources, knowledge, and abilities into devising ideas and solutions that probably would not have been as innovating and productive without conflict.

Unfortunately, the notion of interpersonal skills training for managing conflict has not yet been met with wide social acceptance. For instance, educational systems have done little to integrate such training into their programs. Unfortunately, lack of skills in managing conflict continues to be widespread, and people often find themselves unable to understand how they got involved in escalating conflict and how to handle it. A solution often sought in these situations is intervention by a third party, usually someone with good training in interpersonal skills in conflict management. This section first reviews some of the well-known interpersonal skills for managing conflict. Then, it describes the basic principles and forms of third-party intervention.

4.1. Interpersonal Skills in Managing Conflict

Six basic interpersonal skills for managing conflict have been identified in the past. The first five are viewed as antecedents of the sixth. These basic interpersonal skills must be well acquired and put to effective use before any attempt is made to proceed on to learning how to practice the sixth, that is, problem solving. The latter is the most critical interpersonal skill involved in conflict resolution and, thus, is presented in more detail than the others.

4.1.1. Active Listening

This skill involves listening carefully to what the other person has to say and reflecting on the points he or she is trying to make. Feedback should be provided about what is understood. Points perceived to have ambiguous or double meanings should be clarified with the other by questions that do not include answers and do not assume interpretations.

4.1.2. Empathy

This skill refers to trying to see things from the perspective of the other person. It involves attempting to identify with expectations, fears, and concerns that seem to underlie the other's arguments and responses.

4.1.3. Assertion

This skill suggests that one's point of view, thoughts, feelings, and pursued outcomes are expressed clearly in concrete terms.

4.1.4. Avoiding Dispositional Attributions

Attributing events to the other person's personality, motives, and/or intentions should be avoided at all costs. Communication should be about tasks and issues, whereas affect expressed should not deviate from personal respect and recognition to the other's right to have claims in the relationship.

4.1.5. Constructive Feedback

This skill refers to identifying and rewarding positive behaviors toward resolution while not failing to mention or ignoring negative behaviors. Reasoning should be provided as to why behaviors are positive or negative on the basis of their consequences to self, other, or both.

4.1.6. Problem Solving

Alternative solutions aimed at satisfying the goals and needs of both parties should be generated. This requires (a) each party developing separately a clear hierarchy of goals and needs, with his or her optimum goals placed on top and followed by the maximum minimum and the minimum of the maximum; (b) discussing proposed solutions, beginning with the target of fulfilling optimum goals for both parties and then, if discussions come to an impasse, discussing on the basis of satisfying maximum minimum goals for both parties and then, if this also fails, further discussing on the basis of minimum maximum goals; (c) examining whether maintaining the relationship and maximizing joint outcomes is indeed preferable to pursuing own optimum goals independently; and (d) repeating Steps b and c until both parties concede to aim for lower than their initial outcomes so that a mutually satisfying solution (referred to as "satisficing") can be reached. If necessary, the hierarchy list of preferred outcomes should be reviewed and restructured.

The conflict resolution process may last a long time and may require the control and manipulation of several parameters. Hence, it is important that disputants monitor their effectiveness in managing conflict. Based on Deutsch's 1994 analysis of conflict resolution training, four criteria captured by the following questions can be used to evaluate whether the preceding skills have been put to work effectively. First, has a cooperative, problem-solving relationship with the other person been initiated, and how well is it maintained? Second, do disputants discuss a reasonable range of alternative and mutually satisfying proposals as a result of problem solving? Third, do disputants seem to gradually adopt an "outsider's" perspective or an analytic perspective to the conflict? Fourth, are the various forms of conflict expression kept to a minimum during the process?

Influential applied research, such as Johnson and Johnson's 1994 research on constructive conflict in schools, has consistently documented the benefits of systematic interpersonal skills training in interpersonal educational as well as in other settings. However, as noted previously, educational systems have not yet integrated such training into their programs. Unfortunately, rather than preparing to personally handle the relatively frequent, inevitable, and potentially harmful life events of conflict, people seem to prefer resorting to the

intervention of a third party after conflict has already escalated and negative consequences are already difficult to bear. Although such persistent conflict may have caused a lot of damage or even irreparable harm, it is usually never too late to seek the help of a specialist.

4.2. Third-Party Intervention

A third (independent) party can often be invited by the disputants to resolve conflict. This is more likely when disputants are motivated to maintain their relationship despite its temporary negative outcomes or when they intend to exit the relationship with the maximum possible gain and minimum possible damage for self and other. Third-party intervention may also be offered when conflict has damaging consequences for others indirectly related to the disputants and when the functionality of their social setting is threatened. For instance, in the earlier example, negative consequences had resulted from the conflict between the two psychologists for the functionality of group multispecialist meetings, the work ambience, and the welfare of patients in the small health care institution. In this case, the chief administrator of the institution might have intervened in a number of informal ways. The administrator could have advised the psychologists individually to behave "professionally," to avoid hostility, and to focus on patient care, or the administrator could have facilitated a constructive dialogue between the two. Alternatively, the administrator could have arranged so that the two psychologists never interact or interfere in each other's duties. If these informal approaches had proven to be ineffective, the administrator could have employed a more formal approach such as mediation.

In general, intervention strategies differ in terms of their legalistic and nonlegalistic character as well as the intervening party's interaction with the disputants.

4.2.1. Adjudication

This involves a judge or jury in a court setting. The interaction of these parties with the disputants is strictly formal and follows legally prescribed channels. Disputants are obliged to abide by the court's decision.

4.2.2. Arbitration

This involves little interaction of the intervening party with the disputants. The arbitrator listens to both sides and decides what the solution should be. This approach is also legalistic, to a certain degree, in that norms and prescribed rules must be taken into account in providing a solution. Disputants may have agreed before arbitration to abide by the arbitrator's decision. However, they are not obliged to do so.

4.2.3. Fact Finding

This involves even less interaction with the disputants. The third party's role is to gather information so as to arrive at an independent judgment regarding the dispute. This strategy could also be said to have a legalistic character in the sense described in the preceding subsection (on arbitration). The input of the "fact finder" is only advisory.

Conciliation and mediation, described next, are nonlegalistic processes and tend to be more psychological in nature.

4.2.4. Conciliation

This involves a lot of interaction with the disputants, albeit limited to opening and maintaining communication channels and exchange of information. The conciliator does not get actively involved in disputants' discussions and agreements.

4.2.5. Mediation

This involves an independent third party who must actively interact with the disputants and work with them in identifying the source of conflict and in reaching a mutually satisfying agreement. Mediation is largely preferred to other forms of third-party intervention in many types of interpersonal relationships. This may be due to its relatively lower cost to other, particularly legalistic, conflict resolution procedures or, more important, to the willingness and commitment of anyone choosing mediation to actually resolve conflict. The latter may, in part, account for the finding that mediation yields more effective, lasting, and satisfactory resolutions for disputants. It is also noteworthy that even when conflict is not resolved in a satisfactory way, the mediation process per se tends to improve disputant relations both within and outside the mediation framework.

Psychological research has dealt more with mediation than with any of the other intervention processes, focusing on understanding its mechanisms and its effectiveness. Following is a succinct description of the basic features of this procedure.

The mediator must be a person who is trained in interpersonal skills and has a good understanding of the theoretical underpinnings of conflict and experience in monitoring conflict. Mediation usually starts with preliminary contacts between the mediator and each disputant. The aim of these contacts is twofold: (a) to evaluate disputants' willingness and commitment to resolving conflict and (b) to identify the fundamental source of conflict and of minor issues causing friction. Setting the rules of subsequent interactions follows this stage. The mediator explicates the procedure, delineates his or her role as a neutral moderator, and states emphatically the modes of communication that will and will not be allowed during interaction, usually giving specific examples. For instance, interruptions, dispositional attributions (e.g., name calling), and generalizations and interpretations of the other's arguments should be avoided. Mediation proper then begins. Each disputant is asked to present his or her view of the conflict situation within a specific time limit. The mediator makes sure that no one is interrupted and that the rules are followed. Once both disputants have presented their views, the mediator may summarize each position and asks disputants for further clarifications. This is followed by the problem-solving phase, during which the two disputants and the mediator all collaborate in reviewing (a) each party's interests, goals, and desirables outcomes; (b) each party's sources of interference and nuisance; and (c) as many feasible and desirable solutions to conflict as is possible. Common discussion is followed by separate discussion with each disputant, during which the same issues are reviewed with an emphasis on discussing as many feasible and desirable solutions to the conflict as is possible. During these contacts, the mediator must reevaluate commitment to conflict resolution and exclude the possibility that mediation is being used by either disputant as a strategy to attain a hidden agenda. Furthermore, points of resistance, reluctance, and apprehension regarding the constructive process must be identified, and their causes must be discussed. If this stage yields a range of mutually acceptable alternatives, an agreement can be reached. Agreements are recorded, typed, and signed by both parties. Language used in documents should be specific, concrete, and unambiguous. Disputants are expected to honor their signatures. If the problem-solving stage does not produce a mutually acceptable solution, agreement cannot be reached and the entire procedure must be repeated. The stages described do not necessarily follow one another. An earlier stage might need to be rerun at any point during the mediation procedure. Thus, mediation can be painstaking and time-consuming. Nevertheless, as has often been observed in relevant research, the procedure per se may improve the relationship of the conflicting parties, providing them with a thrust to proceed on to resolving conflict on their own.

5. CONCLUSION

Conflict can undermine the integrity of a relationship and generate negative consequences for the relationship members. However, such detrimental outcomes are due to uncontrollable escalation of conflict and not to conflict per se. In fact, conflict is an integral part of all interpersonal relationships. It is a product of the interdependence that is necessarily developed between two (or more) parties so that they can both (or all) enjoy valued outcomes that could not be enjoyed to the same degree, or with the same ease and quality, outside the specific relationship. Conflict can be managed so as to yield outcomes constructive for the relationship and positive for its individual members. Relationship members should be on guard for uncontrollable conflict escalation, which is usually due to (a) actual or perceived interference of the other with the attainment of own goals, (b) actual or perceived relative deprivation of own outcomes with respect to a commonly accepted standard and the standard attained by the significant other in the relationship, (c) perceiving interference or relative deprivation as a threat to self-esteem, (d) focusing on enhancing self-esteem rather than on specific conflict issues, (e) maximizing own outcomes unilaterally by ignoring the other's or at the expense of the other's outcomes, (f) manifesting and retaliating conflict behaviors, and/or (g) making dispositional attributions about conflict behaviors.

Training in interpersonal skills, such as active listening, empathy, assertion, constructive feedback, and problem solving, can deescalate conflict and turn it into a constructive process. Mediation by a third party is often required when people lack conflict resolution skills. The mediator attempts to facilitate interactions between disputants so that they can negotiate for themselves. Mediation requires analyzing facts, goals, outcomes, and conflict behaviors as reported by each party in individual and common sessions with the mediator. As a constructive process, conflict can regulate self-interest in a relationship. Moreover, it can maintain the relationship's productivity of positive outcomes to a maximum level for the members involved. It can work to clarify needs and goals and to delineate the

role of individual members. It can focus disputants on mutually satisficing solutions of objective conflict differences, moving them away from the persistent and futile effort to restore self-esteem by means of conflict behavior. Last, but not least, it can disentangle individuals from an unproductive relationship that perpetuates their negative outcomes by precipitating exit.

See Also the Following Articles

Interpersonal Behavior and Culture ■ Interpersonal Perception

Further Reading

Deutsch, M. (1994). Constructive conflict resolution: Principles, training, and research. *Journal of Social Issues*, *25*(2), 7–41.

Holmes, J. G., & Murray, S. L. (1996). Conflict in close relationships. In E. T. Higgins, & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 622–654). New York: Guilford.

Johnson, D. W., & Johnson, R. T. (1994). Constructive conflict in the schools. *Journal of Social Issues*, *50*(1). 117–138.

Keashly, L. (1997). Conflict and conflict management. In S. W. Sadava, & D. R. McCreary (Eds.), *Applied social psychology* (pp. 248–273). Upper Saddle River, NJ: Prentice Hall.

Kelley, H. H. (1979). *Personal relationships: Their structure and processes*. Hillsdale, NJ: Lawrence Erlbaum.

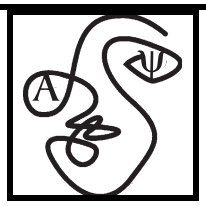
Kressel, K., & Pruit, D. G., & Associates (1989). *Mediation research*. San Francisco: Jossey–Bass.

Lewin, K. (1948). The background of conflict in marriage. In G. W. Lewin (Ed.), *Resolving social conflicts: Selected papers in group dynamics* (pp. 84–102). New York: Harper.

Pruit, D. G. (1998). Social conflict. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (Vol. 2, pp. 470–503). New York: McGraw–Hill.

Rusbult, C. E., & Van Lange, P. A. M. (1996). Interdependence processes. In E. T. Higgins, & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 564–596). New York: Guilford.

Wilmot, W. W., & Hocker, J. L. (2001). *Interpersonal conflict*. New York: McGraw–Hill.



Interpersonal Perception

David A. Kenny
University of Connectinut, Storrs, Connectinut, USA

1. Types of Perceptions
2. Relations among Perceptions
3. Statistical Analysis
Further Reading

GLOSSARY

meta-accuracy How well a person knows how she or he is viewed by others.
metaperceptions The way in which people believe others perceive them.
reciprocity The degree to which people view each other the same (or differently).
self-perceptions How individuals perceive themselves.
self–other agreement How much a person's self-perception is similar to the way in which others see him or her.
target accuracy How well a person really knows another person.

People, whether strangers or marriage partners, form perceptions of one another, and these perceptions have played an important role in social science research. For instance, leadership style was measured by Fiedler using an interpersonal perception measure called LPC, and Newcomb argued that the co-orientation of interpersonal perceptions is key to understanding the formation of friendships. This article considers the types of interpersonal perceptions, the relations among them, and their statistical analysis.

1. TYPES OF PERCEPTIONS

There are three major types of perceptions: perceptions of others, metaperceptions, and self-perceptions. Very often, researchers studying groups and relationships ask people to make judgments of one another. For instance, they may ask people in a work group how hard-working and friendly the other members of the group are. Sometimes, researchers ask observers or outsiders to form perceptions of group and dyad members. Thus, much of what is called an interpersonal "perception" is not so much a perception as an inference made about another person. Nonetheless, the term "interpersonal perception" continues to be used.

Besides asking for perceptions of others, researchers may ask people how they think others view them. These perceptions of perceptions are called metaperceptions or reflected appraisals. Finally, they may ask people how they perceive themselves (called self-perceptions). Although self-perceptions are not themselves interpersonal, they are thought to be closely tied to interpersonal perceptions.

In 1966, Laing and colleagues developed a simple notation for such perceptions. The perception that Person A has of Person B is denoted as A(B), the metaperception is denoted as A(B(A)), and the self-perception is denoted as A(A).

Other more complicated types of interpersonal perceptions are relatively less studied. Laing and colleagues discussed meta-metaperceptions (e.g., What does Person A think that Person B thinks that Person A

©2004 Elsevier Inc.
All rights reserved.

thinks?) and even meta-meta-metaperceptions. Others have studied triadic perceptions (e.g., What does Person A think that Person B thinks of Person C?).

The content of these perceptions is quite varied and is limited only by the imagination of the investigator. Among the topics studied are personality (e.g., Does Person A think that Person B is friendly?), liking (e.g., Does Person A think that Person B likes Person A?), interests and attitudes (e.g., Does Person A think that Person B is politically conservative?), and emotional states (e.g., Does Person A think that Person B is currently angry?). In general, researchers ask perceivers to make judgments about several variables. For instance, a person might be asked whether his or her roommate is friendly, intelligent, lazy, and/or organized.

2. RELATIONS AMONG PERCEPTIONS

Very often, researchers are interested in the relationship among different interpersonal perceptions. In 1994, Kenny, extending the pioneering work of Laing and colleagues, defined a series of questions that relate interpersonal perceptions to each other. Among the most common questions are the following:

- *Consensus:* Do two people agree in their perception of a common target?
- *Reciprocity:* Do two people view each other in the same or opposite ways?
- *Target accuracy:* Does a person really know what another person is actually like?
- *Meta-accuracy:* Does a person know how he or she is viewed by another person?
- *Self–other agreement:* Is a person's self-perception similar to how another person views him or her?

Very often, researchers and theorists use these relations among interpersonal perceptions. For instance, a key tenet of symbolic interactionism is that self–other agreement between a person and a significant other is due to the meta-accuracy of the person knowing how the significant other views him or her and then the incorporation of that view into the self-concept.

3. STATISTICAL ANALYSIS

Research in interpersonal perception can be either ideographic or nomothetic. In ideographic work, the focus is on variation of relationships between perceptions by persons or relationships. An example of an ideographic question is that of whether married couples who know their partners better are more satisfied in their relationships. Most researchers who study interpersonal perception are interested in the ideographic level of analysis.

Alternatively, the focus might be nomothetic or general laws that are true of most persons. For instance, research might ask how much meta-accuracy there is. One advantage of a nomothetic analysis is that it is possible to decompose effects at the level of the individual and the dyad using the Social Relations Model. For instance, one can determine whether meta-accuracy is true of persons across relationships (e.g., Do people know how they are generally seen across relationships?) and whether meta-accuracy is true of particular relationships (e.g., Do people know how they are uniquely seen by particular others?).

There are several difficulties that arise in the analysis of interpersonal perceptions. For instance, researchers sometimes compute the absolute difference or discrepancy between two interpersonal perceptions. Following Cronbach's 1955 work, such scores lead to several interpretive issues. Other issues arise in the analysis of data from interpersonal perception. Despite difficult methodological and statistical issues, interpersonal perception and the relations between perceptions will continue to be important in social science theory and research.

See Also the Following Articles

Perception and Culture ■ Reciprocity Norm

Further Reading

Cronbach, L. J. (1955). Processes affecting scores on "understanding of others" and "assumed similarity." *Psychological Bulletin*, *52*, 177–193.

Fiedler, F. E. (1964). A contingency model of leadership effectiveness. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 1). New York: Academic Press.

Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis*. New York: Guilford.

Kenny, D. A., & DePaulo, B. M. (1993). Do people know how others view them? An empirical and theoretical account. *Psychological Bulletin*, *114*, 145–161.

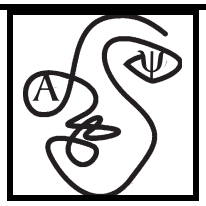
Kenny, D. A., & La Voie, L. (1984). The social relations model. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 18, pp. 142–182). San Diego: Academic Press.

Kinch, J. W. (1963). A formalized theory of self-concept. *American Journal of Sociology*, *69*, 481–486.

Laing, R. D., Phillipson, H., & Lee, A. R. (1966). *Interpersonal perception: A theory and a method of research*. New York: Springer.

McLeod, J., & Chaffee, S. H. (1966). Interpersonal approaches to communication research. *American Behavioral Scientist*, *16*, 467–500.

Newcomb, T. M. (1953). An approach to the study of communicative acts. *Psychological Review*, *60*, 393–404.



Interrogation and Interviewing

Aldert Vrij
University of Portsmouth, Portsmouth, United Kingdom

1. Introduction
2. The Problem: The Interviewee Does Not Provide Enough Information
3. The Problem with Further Questioning: Suggestibility
4. The Solution: Questioning Techniques to Minimize Suggestibility
5. The Solution: Interview Techniques to Enhance Information Recall—The Cognitive Interview
6. The Problem: Persuading Reluctant Interviewees to Talk
7. False Confessions
8. Detecting Deceit
9. Common Pitfalls in Detecting Deceit
10. Techniques to Improve People's Ability to Detect Deceit
11. Conclusion
Further Reading

GLOSSARY

cognitive interview Forensic tool composed of a series of memory retrieval techniques designed to increase the amount of information that can be obtained from a witness.

ethical interviewing A police interview technique, based on three principles, designed to encourage a suspect to tell his or her side of the story; the use of deceit is forbidden, and psychological pressure is not used.

Reid technique A police interrogation technique, consisting of nine steps, designed to encourage reluctant suspects to talk during police interviews; the technique allows interrogators to use deceit and to build up psychological pressure.

suggestibility The act or process of impressing something (e.g., an idea, an attitude, a desired action) on the mind of another.

The purpose of a police interview or interrogation is to obtain information about a possible crime. Effective interviewing is difficult because interviewers risk eliciting inaccurate information that they may misconstrue as accurate.

1. INTRODUCTION

The purpose of a police interview with an alleged victim, witness, or suspect is to obtain information about a possible crime. During this information-gathering process, interviewers should attempt to obtain statements that are complete (i.e., include all the relevant information that the interviewee knows) and reliable (i.e., contain only accurate information). Those two aims are not always compatible. Certain interview techniques readily result in more information, but the information is often less reliable. Psychologists have developed interview techniques that increase the likelihood of obtaining both accurate and reliable information. However, these techniques might not work with interviewees who are reluctant to talk. For this reason, other more persuasive interview techniques have been developed, especially for interviewing uncooperative suspects. Those techniques have been criticized extensively by psychologists. A major problem is that they may induce false confessions, that is, suspects admitting to having participated in crimes when in fact they did not actually participate in those crimes. Obtaining unreliable information would be less problematic if

©2004 Elsevier Inc.
All rights reserved.

police officers could distinguish reliable information from unreliable information.

2. THE PROBLEM: THE INTERVIEWEE DOES NOT PROVIDE ENOUGH INFORMATION

Obtaining information from witnesses, victims, and suspects is critical for effective police investigations. However, the answers to questions such as "Tell me exactly, and in as much detail as possible, what you witnessed" and "Tell me exactly, and in as much detail as possible, what happened to you" are typically incomplete. In nearly all cases, interviewees initially report less information than they actually hold in their memories. Routinely, interviewers want more information than is initially provided. It is not surprising that guilty suspects fail to report all of the information that they are able to recall. Why would they volunteer information that will indisputably link them with their crimes? However, witnesses and victims also rarely volunteer all of the information they know. There are several reasons for this. First, it is a matter of ignorance. Police interviews require witnesses and victims to describe people, objects, and events in more detail than is customary in casual conversation. Often, interviewees do not realize that the police want them to disclose all of the information they know rather than just the information that the interviewees believe is relevant. Second, information is often lacking or incomplete due to cognitive reasons. For example, young children's first accounts of an event are particularly incomplete. Cognitive abilities and command of language develop throughout childhood, making it gradually more feasible to offer detailed accounts of what was witnessed. Older children and adults use more complex and successful retrieval strategies than do younger children, increasing the amount of information they can retrieve independently. People cannot verbally recall experiences that occurred before they learned to speak. Therefore, they cannot verbally relate events that they experienced before 2 years of age. Third, social factors contribute to incomplete accounts. For example, socially anxious people who feel discomfort in the presence of strangers spontaneously volunteer less information. Other people may be too embarrassed to talk freely about certain events that they have experienced or witnessed. For all of these reasons, interviewers must ask further specific questions to learn more about events. However, by asking a series of specific follow-up questions, interviewers run the risk that the questions selected will reflect what the interviewers believe happened rather than what the interviewees actually witnessed or experienced. This risk is most pronounced when the interviewers' questions are suggestive.

3. THE PROBLEM WITH FURTHER QUESTIONING: SUGGESTIBILITY

Despite considerable controversy concerning the extent to which children and adults are susceptible to suggestion in forensic contexts, three conclusions can be drawn. First, there are individual differences in suggestibility. Second, there are age differences in suggestibility. Younger children are more suggestible than older children, and children are more suggestible than adults. Third, the manner in which an interview is conducted can affect suggestibility as well as the quality and extent of information elicited.

4. THE SOLUTION: QUESTIONING TECHNIQUES TO MINIMIZE SUGGESTIBILITY

Research on questioning witnesses and victims in a forensic context has focused on ways in which to avoid suggestibility. Perhaps because children are particularly sensitive to suggestibility, much research has centered on interviewing children. This research shows that open-ended questions (e.g., "Tell me what happened") are best for gathering information because the interviewee has the opportunity to give an open unrestricted answer and to control the flow of information. This type of question produces responses that are longer and richer in relevant details than are responses to other types of questions. Most "wh–" questions (e.g., what, where, when, why, who) can be classified as open-ended questions. These questions should be asked at the beginning of the interview. Facilitative responses by the interviewer, such as "okay" and "mm," encourage the witness to continue the account. Focused questions can be asked later in the interview. These are questions that focus the witness's attention on details or aspects of the event that the witness described previously. For example, "Tell me what the man looked like" is an example of a focused question after an interviewee mentioned a man but did not describe his appearance. Leading questions (e.g., "Was the man Black?") are problematic because the interviewee may respond by

simply repeating the information conveyed in the question rather than by relying on his or her memory. Option-posing questions, or questions that offer alternatives (e.g., "Was the man Black or White?"), are preferable in that respect because they are less leading. However, there is still the chance that the interviewee's preferred option might not be incorporated into the alternatives posed by the interviewer. An open-ended question (e.g., "What was the color of the man's skin?") is a less suggestive question format. Where leading questions are asked (e.g., "Did he touch you?"), they should be followed by open-ended questions designed to elicit further information about the topic (e.g., "Tell me everything about that").

Few practitioners apply these principles in forensic interviews. For example, in 2002, Westcott and Brace reported an analysis of 119 interviews with children by police officers and social workers in England and Wales. The analysis revealed that open questions were rarely asked (6%) and that focused questions were most common (47%). Questions that included alternatives occurred more frequently (29%) than did facilitative responses (13%), and some questions were leading (5%). Similar studies conducted in Sweden, Israel, and the United States provided similar results.

5. THE SOLUTION: INTERVIEW TECHNIQUES TO ENHANCE INFORMATION RECALL—THE COGNITIVE INTERVIEW

Another line of research has explored whether the quality and amount of information provided by cooperative interviewees can be improved by certain interview techniques. One such technique, the cognitive interview, has been adopted by police forces around the world, including Australia, Germany, the United Kingdom, and the United States.

The cognitive interview, developed by psychologists, has three core elements: social dynamics, memory/general cognition, and communication. Social dynamics and communication refer to interpersonal communication and ways in which to deal with anxious or inarticulate witnesses. The interviewing police officer has ostensible social authority and more control over the interview than does the witness and, therefore, may dominate the interview. To avoid counterproductive dynamics, when interviewing an cooperative witness, the police officer should explicitly ask the witness to take an active role in the interview.

Officers should communicate with witnesses in the style that the witnesses prefer by adopting witness-compatible questions. Most often, depictions will be verbal, but some witnesses find it easier to describe certain events nonverbally. For example, if the event that a witness experienced involved a tactile element such as brushing against a fabric, the witness might prefer to recall this experience in a similar mode by touching various fabrics.

The memory/general cognition element consists of four components. A "report everything instruction" is given to encourage witnesses to report everything they remember without any editing, even if the details seem trivial or their memory of a particular aspect of the event is partial or incomplete. Memory retrieval is most efficient when the context of the original event is mentally reinstated at the time of recall. Therefore, interviewees are instructed to reconstruct in their minds the environmental features (i.e., location where the event took place) and personal features (i.e., how they felt at the time) of the event. Different retrieval cues may access different aspects of a complex event. Therefore, after witnesses recount the event in their own order during free recall, they should be asked to recall the event using a variety of different temporal orders, for example, in reverse order (i.e., from the end of the event to the beginning of the event). People have a tendency to report events from their own perspective. An instruction to change perspective asks interviewees to recall the event from the perspective of another person who was present (e.g., "Go through the event again and tell me any further information witnessed by that other person you mentioned").

Studies testing the efficacy of the cognitive interview have revealed that this technique elicits more accurate information than do conventional interview techniques. This technique also elicits more inaccurate information in that it increases the absolute number of errors; however, it does not increase the rate of inaccuracy. In other words, the percentage of information that is inaccurate is not higher using the cognitive interview than that obtained using conventional interview techniques.

6. THE PROBLEM: PERSUADING RELUCTANT INTERVIEWEES TO TALK

The stereotypical view, often promoted in police manuals, is that suspects are reluctant to talk and that, therefore, tough interrogation techniques, rather than

interview techniques, are needed to persuade suspects to talk. The assumption that suspects are unwilling to talk is plausible because confessing to a crime dramatically increases the likelihood of a conviction. For example, Dutch research revealed that in cases where the suspects confessed, convictions were nearly inevitable (a 94% conviction rate was found), whereas without a confession, the conviction rate was only 44%. However, the view that suspects are generally uncooperative is not supported by research. Studies in England and Wales revealed that 80% of suspects are cooperative during interviews and that 55 to 60% of suspects confess. (Much research on this topic has been carried out in England and Wales because, by law in those countries, all police interviews with suspects must be audiotaped.) These studies also examined what prompts suspects to confess. The most important factor seems to be the perceived strength of evidence; suspects are more likely to confess if they perceive the evidence against them to be strong. Another influential factor is the seriousness of the crime; suspects are less inclined to confess when they are suspected of committing more serious offenses. Hence, suspects are reluctant to confess when they are suspected of serious crimes and/or the evidence against them is weak.

Several interrogation methods have been proposed to persuade reluctant suspects to talk. Probably the most well known is the "Reid technique" proposed by Inbau and colleagues. The back cover of the latest edition of their manual, published in 2001, proclaims that their technique is "the standard in the field" and that it has been used successfully by "thousands of criminal investigators." The Reid technique is meant to apply to suspects who are likely to be guilty. Interviewers are advised to make preliminary assessment of the likelihood of guilt by observing suspects' verbal and nonverbal responses. For example, according to Inbau and colleagues, verbal responses typical of untruthful suspects include giving delayed, evasive, and/or vague answers; displaying unusually poor memory; making statements against self-interest (e.g., "as crazy as it sounds"); and answering questions with questions (e.g., "Why would I do that?"). Assumed behavioral clues to deception include changing posture, maintaining a retreating posture (e.g., crossing arms in a tight fashion), making grooming gestures, placing a hand over the mouth or eyes when speaking, averting one's gaze, and hiding hands (by sitting on them) or hiding feet (by pulling them under the chair).

6.1. The Reid Nine Steps of Interrogation

The Reid technique strives to break the (guilty) suspect's resistance to confess. It consists of the following nine steps:

1. Engaging in a positive confrontation
2. Developing a theme
3. Handling denials
4. Overcoming objections
5. Retaining the subject's attention
6. Handling the suspect's mood
7. Creating an opportunity to confess
8. Obtaining an oral confession
9. Converting an oral confession into a written one

Step 1 consists of a direct presentation of real or fictional evidence regarding the suspect's involvement in the alleged crime (e.g., "Our investigation shows that you are the one who . . ."). This accusation is followed by a brief pause during which the suspect's behavior is observed closely. The interrogator repeats the accusation. Suspects who fail to make a denial after the second direct confrontation are presumed to be guilty.

Step 2 deals with theme development tailored to emotional versus nonemotional suspects. In the case of an emotional suspect, the interrogator is advised to build rapport by offering the suspect a moral excuse for committing the offense. This can be achieved by rationalizing the situation, suggesting that anyone else in the same situation might have committed the same type of offense, minimizing the moral seriousness of the offense, suggesting a more morally acceptable motivation for the offence, condemning others, and so on.

Suspects who appear to listen attentively to these suggestions are presumed to be guilty. The following example, from Inbau and colleagues, illustrates this technique: "Joe, no woman should be on the street alone at night looking as sexy as she did. . . . It's too much a temptation for any normal man. If she hadn't gone around dressed like that, you wouldn't be in this room now." Critics of the technique label it the minimization approach and describe it as a technique in which the police interrogator tries to lull the suspect into a false sense of security by offering sympathy.

Nonemotional suspects are expected to perceive an interrogation as a "contest of endurance," that is, as "pitting their willpower against the interrogator's persistence." Effective techniques for these cases are (a) to seek an admission of lying about some incidental aspect of the crime so that the interrogator can say

"You lied to me once and you will lie to me again" and (b) to encourage the suspect to associate himself or herself with the scene of the crime, pointing out that all of the evidence leads to the suspect's guilt, that it is futile to resist telling the truth, and so on. Critics call this the maximization approach, a technique in which the interrogator tries to scare and intimidate the suspect into confessing by exaggerating the seriousness of the offense and the magnitude of the charges and, if necessary, by making false claims about the evidence.

Step 3 describes how to respond to denials of guilt. The goal is to curtail the suspect's repetition or elaboration of the denial (using techniques described in Step 2) and is seen as a crucial step because the more frequently a guilty suspect repeats a lie, the harder it becomes for the interrogator to persuade the suspect to tell the truth. The theory is that an innocent suspect typically will persist in denying guilt, whereas a guilty suspect is more likely to allow the interrogator to return to the conversation theme. An innocent suspect is expected to maintain denial, and a guilty suspect is expected to offer reasons why the accusation is wrong (e.g., "I couldn't have done that; I don't own a gun").

Step 4 consists of overcoming the suspect's objections. For example, the interrogator is advised to show understanding and then return to the conversation theme (e.g., "That may be true, but the important thing is . . ."). Once the suspect observes that his or her objections are ineffective, the suspect is expected to become uncertain and display signs of withdrawal. The interrogator is advised to react by retaining the suspect's attention (Step 5), for example, by moving physically closer to the suspect, leaning towards him or her, and maintaining eye contact. Step 6 involves strategies to manage the suspect's passive mood. When the suspect appears to be attentive, the interrogator is directed to focus the suspect's mind on possible reasons for committing the crime. The interrogator should exhibit signs of understanding and sympathy and should urge the suspect to tell the truth. Finally, the interrogator should attempt to elicit a remorseful mood, for example, by pointing out the negative consequences of the crime for the victim. In Step 7, alternative question forms are recommended to offer the suspect the opportunity to provide an explanation or excuse for the crime, and this will make an admission easier to achieve (e.g., "Was this your own idea or did someone talk you into it?"). Whichever alternative a suspect chooses, the net effect of an expressed choice is an admission. Following the selection of an alternative, Step 8 consists of the development of the initial admission into a detailed confession that discloses the circumstances, motives, and details of the crime. Inbau and colleagues advised that the questions asked by the interrogator should be brief, clear, and designed so that they can be answered in a few words. Furthermore, they should not contain harsh or emotionally charged terminology. Step 9 deals with converting an oral confession into a written one. This step is seen as important because the suspect might deny ever making an oral confession.

6.2. Concerns with the Reid Technique

The Reid technique has been criticized extensively by academic researchers. Numerous assumptions presented in the manual, such as "two denials imply guilt" (Step 1), have never been tested empirically. Other assumptions, such as how liars behave and what they say, have been tested but not supported.

By asking (during the confession phase, i.e., Step 8) leading questions that can be answered by the suspect in a few words, the interrogator runs the risk that the confession will reflect more about what the interrogator believes happened than what actually transpired.

In addition, bluffing (i.e., making up evidence) may backfire. For example, if an interviewer lies and tells a burglar that his or her fingerprints have been found at the scene of the crime, a burglar who wore gloves now knows that this statement is a lie and will probably realize that the police do not actually have much evidence against him or her. Inbau and colleagues acknowledged this problem and recommended that this tactic be used only as a "last resort effort." Researchers have demonstrated that North American police detectives can be very creative in devising examples of bluffing such as introducing nonexistent technology (e.g., "The Cobalt Blue test has proven that the fingerprints on the body of the victim belong to you"). Research has revealed that confronting a suspect with false evidence is quite common in the United States and happens in approximately 30% of custodial interviews. Although these sorts of tricks are allowed in the United States, they are illegal in several Western European countries. Police detectives who try to apply the Reid technique in those countries run the risk that subsequent confessions will be dismissed in criminal courts.

The Reid technique of psychologically manipulating people through trickery and deceit may lead to false confessions. One problem with tricks and deceit is that

both guilty and innocent suspects become more willing to confess. There is no strategy that motivates only guilty suspects to talk. As a result, the more successful a strategy is in eliciting confessions from guilty suspects, the more likely it is that this strategy will also produce false confessions from innocent suspects.

6.3. The Solution: Ethical Interviewing

In the United Kingdom, an ethical framework of police interviewing is advocated. The British outcry for, and subsequent legislation of, less oppressive police investigative interview techniques has probably been heavily influenced by cases such as "The Guildford Four" and "The Birmingham Six," well-known cases of British miscarriages of justice. Ethical interviewing is based on three principles: (a) to shift the police service from its traditional reliance in getting a suspect to confess to encourage a search for the truth, (b) to encourage officers to approach an investigation with an open mind, and (c) to encourage officers to be fair. Police training in this ethical framework emphasizes the importance of detailed information gathering at the beginning of the police investigation. It is argued that before the interview starts, the detectives should have been to the scene of crime and, therefore, know how it looks; should know all of the facts of the crime (e.g., evidence obtained at the scene of the crime, statements of witnesses, any peculiarities); and should know the suspect (e.g., background characteristics, family circumstances, possible addictions, possible diseases).

It is further stressed how to use open and closed questions in police interviews. An information-gathering strategy at the beginning of the interview is advocated because this increases the possibility of eliciting an account from the suspect. The use of open questions is generally preferable. They usually elicit longer answers and, therefore, more information. They are an invitation to the suspect to present his or her point of view and will increase the likelihood of the suspect believing that the interviewer is taking him or her seriously. Open questions encourage the suspect to talk and, therefore, facilitate the desired format of a police interview; that is, the suspect talks and the interviewer listens and asks for clarifications. An interviewer who is prepared to listen is more likely to be liked by the suspect, and this might make the suspect more willing to talk. Closed questions are also useful in interviews, particularly to obtain short factual answers on specific points. However, these questions should be used sparingly at other times and rarely during the early stages of the interview.

Several strategic tactics are offered, including the "blocking escape routes" method, which is used to strengthen the evidence in a case. Evidence is sometimes thin and open to multiple interpretations. Presenting this evidence at an early stage of the interview may give the suspect the opportunity to "escape" by providing alternative explanations. The blocking escape routes method is designed to prevent this from happening. For example, suppose that the suspect's car was noticed near the scene of the crime just after the crime took place. This might be a link, but the link is not strong. If the suspect is confronted with this piece of evidence at an early stage of the interview, the suspect might, for example, say that he or she used the car to go to a store. The evidence is much stronger if, before the evidence is presented, the suspect has told the interviewer (after being asked about this) that he or she did not use the car on that particular day, that he or she never lent the car to anyone else, and that nobody else has the keys to the car. After the escape routes have been blocked, the suspect could be confronted with his or her own statements and evidence (e.g., "You told me that you are the only one who uses that car, right? Well, a high-speed camera provided evidence that you drove with your car at high speed near the scene of the crime just after the crime took place. Could you explain that?").

At first sight, the ethical interviewing approach appears to be "softer" than the Reid technique; however, this is not necessarily the case. Although the use of trickery and deceit is forbidden in an ethical interviewing framework, a "tough" interview is allowed. It might be that suspects who decided not to talk during their interviews will remain silent without the use of trickery and deceit, and in such cases the ethical approach will not be successful for the interviewing of uncooperative suspects. This poses a problem, especially if these uncooperative suspects are believed to have committed serious crimes. Ethical methods designed to make silent suspects more willing to cooperate have yet to be developed. The difficulty is that any such methods are likely to make both guilty and innocent suspects more willing to cooperate, and the latter could lead to false confessions. Inbau and colleagues acknowledged the problem of false confessions and offered the following solution:

> A guideline that an interrogator may use in any case situation where he may be in doubt as to the

> permissibility of any particular type of trickery and deceit, is to ask himself the following question: "Is what I am about to do, or say, apt to make an innocent person confess?" If the answer is "no," the interrogator should go ahead and do or say what was contemplated. On the other hand, if the answer is "yes," the interrogator should refrain from doing or saying what he had in mind.

The obvious problem, not addressed by Inbau and colleagues, is that the police detective never knows for sure whether he or she made a correct judgment.

7. FALSE CONFESSIONS

How often false confessions occur is unknown, but estimates range from 35 to 600 convictions on the basis of false confessions per year in the United States alone. It is often difficult to establish whether a confession is false. Police coercion during the interview is not determinative because a confession obtained using an oppressive interview style could be truthful. The retraction of a confession is another potential indicator. On the one hand, a confession may be truthful even though the suspect retracts the confession after the interview; on the other hand, suspects convicted on the basis of confessions do not necessarily appeal their convictions but still could have confessed falsely. Finally, one might consider the accuracy of the confession. On the one hand, a confession in which accuracy cannot be established beyond doubt (e.g., because independent case facts are not available) is not necessarily false; on the other hand, innocent suspects may at times give accurate accounts of what has happened because police detectives provided case information to the suspects, who subsequently incorporated the information into their statements.

Why suspects make false confessions is difficult to determine, but several factors are often present, for example, a combination of tricks employed by the police, a susceptible suspect, and interviewers determined to secure a confession.

7.1. Types of False Confessions

Psychologists distinguish among three psychologically distinct types of false confessions: voluntary false confessions, coerced compliant false confessions, and coerced internalized false confessions.

7.1.1. Voluntary False Confessions

Voluntary false confessions are given without any external pressure from the police. Most often, people voluntarily go to the police station and inform the police that they have committed crimes reported on television or in newspapers. Voluntary false confessions occur for several reasons. One reason is a "morbid desire for notoriety," that is, a pathological need to become infamous and to enhance self-esteem, even if it entails the prospect of imprisonment. One serial false confessor, Henry Lee Lucas, is estimated to have confessed to more than 600 murders during the early 1980s. When asked why he confessed to all of these crimes, Lucas answered that prior to his arrests, he had no friends and nobody listened to him or took any interest in him; however, once he began to make false confessions, all that changed and he now had many friends and enjoyed his celebrity status. Other people may voluntarily falsely confess because they are unable to distinguish facts from fantasy. Schizophrenic people, in particular, are prone to this type of confession. Yet another reason why people may voluntarily falsely confess is that they want to protect the real criminals. For example, a father might falsely confess to a crime that he knows or believes was actually committed by his son. Finally, people may voluntarily falsely confess because they see no possible way in which to prove their innocence and hope that confessing will secure reduced punishment.

7.1.2. Coerced Compliant Confessions

Coerced compliant confessions occur when, as a result of police interviews, suspects confess to acts they know are untrue. This type of confession results from the pressures of police interviews and from the strategies used by the police in the interviews. As a result, suspects believe that the benefits of confessing outweigh the costs, and so they confess to escape from the police interviews, which they consider stressful and intolerable. Alternatively, they confess due to police promises of rewards following their confessions, or they confess due to a combination of the two.

7.1.3. Coerced Internalized False Confessions

Coerced internalized false confessions occur when people come to believe, during police interviewing, that they have committed the crimes of which they are accused, even though they have no actual memories of committing the crimes. One vivid example is

the case of Tom Sawyer, a socially anxious man and former alcoholic who came to believe that he committed a murder that he in fact never committed. In trying to engage Sawyer in conversation about the crime, the detectives asked him to help them create a scenario describing how the murder might have happened. Sawyer, who loved to watch detective shows on television, was eager to help and joined in. The police let Sawyer outline several scenarios and accused him of having committed the murder. The police claimed that Sawyer knew nine facts that only the killer could have known. Analysis of the interrogation transcripts showed that all of the crucial information was introduced into the interrogation by the police. Following the accusation, Sawyer strongly denied his guilt. The police obtained fingerprint and hair samples from Sawyer and suggested a polygraph examination. Sawyer believed that the polygraph would prove his innocence and agreed to the examination. After the test, the examiner told Sawyer that the test proved he was lying. (Subsequent rescoring of the test by a polygraph expert revealed that the test outcome was inconclusive.) Once told that the polygraph showed him to be lying, Sawyer's confidence began to erode. He was no longer able to express firm denials of guilt, and he could say only that he still did not believe he had done it. His main defense against the police demands to confess was his lack of memory of having committed the crime. The police replied that he was blocking his memory of the crime and that he had a blackout, just as he often did when he had been drinking. Sawyer still refused to fully accept that he had committed the crime, and he hoped that the fingerprint and hair tests would prove his innocence. The detectives decided to lie and told Sawyer that his hair samples matched hairs found on the victim's body. On receipt of this information, Sawyer's resistance collapsed, and he agreed that "all the evidence was in" and that he must have committed the crime.

The fact that police deception leads to false confessions might appear to be surprising. However, police suspects are often not average people. Typically, their intelligence quotients (IQs) are lower than average. A sample of 160 suspects in England had an average IQ of 82. People with lower IQs are more suggestible. Sawyer was also socially anxious and a former alcoholic. The combination of police deception and the vulnerability of the suspect may have led to the false confession.

Internalized false confessions do not imply that suspects become convinced that they have committed crimes. They occur because suspects form opinions that it is more likely than not that they are guilty. They typically have no memory of committing the crimes, but tactics used by the police diminish the suspects' confidence in their memories, make them less certain about their innocence, and make them consider whether it is possible that they had committed the crimes.

7.2. How to Detect False Confessions

The best way in which to distinguish between false and true confessions is to look for intimate knowledge and impossibilities. Intimate knowledge is present when a suspect provides information that only the guilty person could know (e.g., information about the location of a disposed body). Impossibilities refer to providing evidence that demonstrates conclusively that the person could not have committed the crime. For example, Lucas's confession could not be true because he was more than 1,300 miles from the scene of the crime on the day of the murder. The psychologist who conducted a psychological evaluation of Lucas considered the uncritical willingness to accept his confession as the most fundamental mistake made by the police. However, police officers claim that it is the responsibility of the courts, and not the police, to check the veracity of a confession. Courts tend to attribute impossibilities to errors in confessions rather than considering them as evidence of false confessions.

Obtaining unreliable information, such as false confessions and incorrect answers to leading questions, would be less problematic if police or jurors could distinguish reliable statements from unreliable ones. A distinction should be made between unreliable (inaccurate) information, which is given deliberately and which the presenter himself or herself knows is inaccurate, and inaccurate information, which is given unwittingly (e.g., because the person is mistaken or no longer knows for certain what happened). Virtually no research has been carried out examining whether people can detect these latter fortuitous truth distortions. The few findings that do exist are not encouraging. Considerable research efforts have examined whether observers can distinguish between truths and deliberate truth distortions (i.e., lies).

8. DETECTING DECEIT

Research has demonstrated convincingly that catching liars is a difficult task and that people frequently

err when they attempt to detect deceit. In scientific studies on the detection of deception, observers are typically given videotaped statements by a number of people who are either lying or telling the truth. After each statement, observers (often college undergraduates) are asked to judge whether the statement is truthful or false. In such tasks, guessing whether someone is lying has a 50% chance of being correct. A review of 37 lie detection studies, in which undergraduates attempted to detect truths and lies told by people (adults) with whom they were unfamiliar, revealed a total accuracy rate (percentage of correct answers) of 56.6%, barely above the level of chance. People are shown to be capable of detecting truths to some extent (i.e., 67% accuracy in correctly judging that someone is telling the truth) but are particularly poor at detecting lies (i.e., 44% accuracy in correctly judging that someone is lying). In fact, 44% is below the level of chance. In other words, people would be more accurate at detecting lies if they simply guessed. The superior accuracy rate for truthful messages is the consequence of the truth bias. People are more likely to expect messages to be truthful than to be deceptive, and as a result, truthful messages are identified with more accuracy than are deceptive ones.

In other studies, people attempted to detect children's truths and lies using the same methodology except that the liars and truth tellers were children rather than adults. That research revealed that detecting children's lies is difficult as well, with accuracy rates typically in the 50 to 65% range.

Undergraduates are not habitually called on to detect deception, whereas police officers and customs officers have more experience in interviewing people. However, studies in which professionals such as police officers attempted to detect truths and lies told by adults revealed accuracy rates similar to those found in studies with undergraduates, mostly in the 45 to 60% range. This research suggests that professionals are no better than laypersons at detecting deceit. However, some professionals (e.g., members of the Secret Service, U.S. federal officers with experience and a special interest in deception and demeanor, sheriffs identified by their departments as outstanding interrogators) were better than others. Professionals are less susceptible to the truth bias than are lay observers, possibly because their jobs make them more suspicious and, hence, more likely to conclude that a person is lying.

9. COMMON PITFALLS IN DETECTING DECEIT

There are numerous reasons why people are generally poor at detecting deceit. Above all, lie detection is a notoriously difficult task because there are no typical behaviors associated with deception. In other words, a cue such as Pinocchio's growing nose does not exist. Despite this, people, laypersons, and professionals have strong views about how liars behave. Interestingly, their views do not differ substantially. In particular, people believe that liars display clear signs of nervousness such as averting one's gaze and making grooming gestures. In reality, however, gaze behavior and grooming gestures are unrelated to deception. If there is a relationship between movements and deception, it is more likely to be the converse, that is, that liars tend to decrease their movements rather than increase them. There are several explanations for this. First, it might be the result of cognitive load. Sometimes, liars find it difficult to lie because they must think of plausible answers, avoid contradicting themselves, and tell lies that are consistent with everything the observers know or might find out. Moreover, liars must remember what they have said so that they can say the same things again when asked to repeat their stories. This might be more difficult than truth telling, especially when liars are unprepared and must make up their stories instantly. Increased cognitive load results in a neglect of body language, reducing overall animation. Second, liars might worry that several cues will give their lies away and, therefore, try to suppress such signs and to make convincing impressions. However, it is not easy to suppress nervousness effectively, to mask evidence of concentration, to know how they normally respond so as to make honest and convincing impressions, and to display the intended responses. It may well be the case that, when trying to appear truthful, liars "overcontrol" their behavior and exhibit a pattern of behaving that appears planned, rehearsed, and lacking in spontaneity. For example, liars may believe that making lots of movements will give their lies away and, therefore, will move very deliberately and tend to avoid any movements that are not strictly essential. This will result in an unusual degree of rigidity and inhibition because people do normally make movements that are not essential. Another possible cue that may result from inadequate control of behavior is that performances might look flat due to a lack of involvement. For example, someone who applies for a job because the money appeals to him or her might not be enthusiastic when the

person talks about the job during the selection interview, and this in turn might result in a lack of animation.

Police officers' incorrect beliefs about cues to deception, such as liars' gaze aversion and grooming gestures, are particularly harmful for suspects who exhibit these patterns of behavior naturally. There is evidence that people from Western cultures exhibit this pattern of behavior less frequently than do people from other ethnic origins such as Black/African Americans and people from Mediterranean countries (e.g., Turkey, Morocco). Some behaviors are culturally defined. For example, looking an authority figure straight in the eyes is considered to be impolite in many cultures, but the opposite is true for Caucasian people living in, for example, North America or Western Europe. For them, averting gaze from a conversation partner is suspicious. Therefore, when a White police officer interviews a suspect from a different ethnic background, cross-cultural nonverbal communication errors may arise. Specifically, natural truthful behavior displayed by a member of a different ethnic group (e.g., gaze aversion) may be incorrectly interpreted as a sign of deceit.

Another pitfall in lie detection is that experiencing and displaying negative emotions is not the exclusive domain of liars. Truth tellers might experience similar negative emotions, albeit for different reasons. For example, lying suspects might be nervous during police investigations because they fear that they will be caught in their lies. Innocent suspects might also be nervous because they fear that they will not be believed by the police detectives when they maintain their innocence. Therefore, lie detectors should be careful in interpreting signs of fear as signs of deceit, and they should avoid making the mistake committed by Othello in Shakespeare's play. Othello falsely accused Desdemona (his wife) of infidelity. He told her to confess because he was going to kill her for her treachery. Desdemona asked for Cassio (her alleged lover) to be called so that he could testify to her innocence. Othello told her that he had already murdered Cassio. Realizing that she could not prove her innocence, Desdemona reacted with an emotional outburst. Othello misinterpreted this outburst as a sign of her infidelity.

10. TECHNIQUES TO IMPROVE PEOPLE'S ABILITY TO DETECT DECEIT

To date, a truly accurate lie detection technique has yet not been developed. All of the proposed techniques produce frequent errors. Police training manuals often emphasize that lie detectors should pay attention to someone's behavior to detect deceit because people are less able to control their behavior than to control their speech. Researchers expressed a similar view for many years. However, more detailed analyses of people's speech, starting in the late 1980s, revealed that there are many more speech-related cues to deceit than was thought previously. For example, compared with stories told by liars, truth tellers' stories sound more plausible and less ambivalent. Liars also tend to tell their stories in chronological order (this happened first, and then this, and then that, etc.), whereas truth tellers sometimes tend to give their accounts in unstructured ways, particularly when talking about emotional events. Sometimes, when people are clearly upset, they report what happened in chaotic and incomprehensible ways. In fact, their stories can be so incomprehensible that listeners must ask them to sit down for a while, to calm down, and to describe again exactly what happened, starting from the beginning. This unstructured production effect disappears when people have already told their stories a couple of times, or when they have frequently thought about the events, because this results in the telling of stories in a more chronological order. Moreover, truth tellers include more visual details (i.e., describe things they saw) and more auditory details (i.e., describe things they heard) in their accounts. Often, they also tend to repeat what was said literally (e.g., "Then he said, 'How are your parents?" ' rather than "Then he asked me about my parents"). Moreover, truth tellers mention more spatial details (e.g., "He stood behind me," "The book was on the table under the window," "We were in the living room") and more temporal details (e.g., "He switched on the television first and then the video recorder," "The phone rang again approximately 15 minutes later," "He stayed with me for about 3 hours"). Research has demonstrated that observers who concentrate on these speech-related cues typically make better lie detectors than do observers who do not attend to these cues.

Another way in which to improve people's ability to detect deceit is to use a so-called baseline method. There are large individual differences in people's speech and behavior. Some people typically make many movements, whereas others do not. Some people are very talkative, whereas others are not. Because of these individual differences, a given response is often difficult to interpret given that it leaves the observer with a dilemma: Is a particular response a sign of deceit, or is it someone's natural truthful response? Therefore, it might facilitate

lie detection if someone's natural truthful speech style and behavior are known to the observer so that the behavior and speech style under investigation can be compared with their natural truthful speech style and behavior. A deviation from the natural truthful response might indicate that a person is lying. Research has shown that this baseline method works and that people are better lie detectors when they know somebody's natural truthful behavior. However, employing this method is more difficult than people might think initially. Not only do different people behave differently in the same situation (i.e., interindividual differences), but the same person behaves differently in different situations as well (i.e., intraindividual differences). Therefore, in applying the baseline method, it is crucial that only parts of the interview that are truly comparable be compared. Comparing the wrong parts of an interview is a common practice in police interviews, and unfortunately, police manuals sometimes advise police detectives to make comparisons that are inappropriate. Typically, comparisons are made between responses during "small talk" at the beginning of the interview and responses during the actual interrogation. This is an incorrect way in which to employ the technique because small talk and the investigative core of the police interview are fundamentally different. The formal interview matters to suspects because severe negative consequences may follow if they are not believed. In other words, this is a high-stakes situation. The small talk has no such consequences and, therefore, is a low-stakes situation. People tend to behave differently in low-stakes situations than they do in high-stakes situations. Both guilty and innocent suspects tend to change their behavior the moment their formal interviews start. For comparisons to be truly comparable, they must be made within formal interviews. For example, a suspect could be asked to describe in detail his or her activities during a particular day (i.e., the day the crime was committed), and the interviewer could observe the suspect's reactions while he or she gives the description. However, even when correct comparisons are made, the lie detector should be cautious about drawing conclusions. A difference in response between two comparable parts of the interview does not indicate conclusively that someone is lying; it only indicates that this is a possibility. Other explanations for the change in response should be taken into consideration as well. For example, there is some evidence that some behaviors are topic related. ("Topic" is not synonymous with "stakes." During a high-stakes interview, different topics can be discussed.) Analyses of the behavior by Saddam Hussein when he was interviewed by Peter Arnett during the first Gulf War revealed that the supportive arm and hand movements Hussein made were, to some extent, related to the topics he discussed. When discussing Israel, he made specific movements with his left forearm. This behavioral pattern emerged only when he was discussing Israel and Zionism. Therefore, the lie detector should consider whether different response patterns shown during a high-stakes interview might be topic related.

11. CONCLUSION

People, including professionals, have difficulty in discerning lies. It is unlikely that detection of unintentional truth distortions is easier (although this is rarely investigated). Police investigators should question interviewees—witnesses, alleged victims, and suspects—in ways that decrease the likelihood of obtaining inaccurate information, as outlined in this article.

See Also the Following Articles

Ethics and Social Responsibility

Further Reading

Baldwin, J. (1993). Police interview techniques. *British Journal of Criminology*, *33*, 325–352.

DePaulo, B. M., & Friedman, H. S. (1998). Nonverbal communication. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (pp. 3–40). New York: McGraw–Hill.

DePaulo, B. M., Lindsay, J. L., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, *129*, 74–118.

Ekman, P. (2001). *Telling lies: Clues to deceit in the marketplace, politics, and marriage*. New York: Norton. (Originally published 1985)

Fisher, R. P., Brennan, K. H., & McCauley, M. R. (2002). The cognitive interview method to enhance eyewitness recall. In M. L. Eisen, J. A. Quas, & G. S. Goodman (Eds.), *Memory and suggestibility in the forensic interview* (pp. 265–286). Mahwah, NJ: Lawrence Erlbaum.

Gudjonsson, G. H. (2003). *The psychology of interrogations and confessions: A handbook*. Chichester, UK: Wiley.

Inbau, F. E., Reid, J. E., Buckley, J. P., & Jayne, B. C. (2001). *Criminal interrogation and confessions* (4th ed.). Gaithersburg, MD: Aspen.

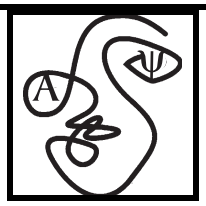
Kassin, S. M. (1997). The psychology of confession evidence. *American Psychologist*, *52*, 221–233.

Vrij, A. (2000). *Detecting lies and deceit: The psychology of lying and its implications for professional practice*. Chichester, UK: Wiley.

Vrij, A. (2003). "We will protect your wife and child, but only if you confess": Police interrogations in England and The Netherlands. In P. J. van Koppen, & S. D. Penrod (Eds.), *Adversarial versus inquisitorial justice: Psychological perspectives on criminal justice systems* (pp. 55–79). New York: Plenum.

Westcott, H., & Brace, N. (2002). Psychological factors in witness evidence and identification. In H. Westcott, & N. Brace (Eds.), *Applying psychology* (pp. 117–178). Milton Keynes, UK: Open University Press.

Williamson, T. (1993). From interrogation to investigative interviewing: Strategic trends in police questioning. *Journal of Community and Applied Social Psychology*, *3*, 89–99.



Intrinsic and Extrinsic Motivation in Sport

Robert J. Vallerand

Université du Québec à Montréal, Montreal, Canada

1. Introduction
2. Definitions and Assessment
3. On the Determinants of Intrinsic and Extrinsic Motivation
4. On Motivational Outcomes
5. Interventions
6. Conclusion
   Further Reading

GLOSSARY

amotivation A lack of purpose and intentionality in one's action, that is, the relative absence of motivation.

extrinsic motivation Engaging in an activity as a means to an end and not for the activity's own sake.

intrinsic motivation Engaging in an activity for itself and for the pleasure and satisfaction derived from participation.

motivation The hypothetical construct that is used to describe the internal and/or external forces that lead to the initiation, direction, intensity, and persistence of behavior.

need for autonomy The desire to be the origin of one's behavior.

need for competence The desire to interact effectively with the environment.

need for relatedness The desire to be connected to significant others in one's interaction with the environment.

recursive bottom-up effect The impact that specific motivations have in a repeated matter over time on more generalized forms of motivation.

self-determination A sense of personal initiative, freedom, and choice in behavior.

self-determined motivation Motivation that is characterized by high levels of intrinsic motivation and identified regulation but low levels of external and introjected regulation and amotivation.

top-down effect The impact of more generalized motivational orientations on more specific motivations.

The purpose of this article is to provide an overview of what is known about intrinsic and extrinsic motivation in sport. After some definitional issues, the article focuses on the assessment of intrinsic and extrinsic motivation. It then proceeds to a discussion on the determinants and outcomes of intrinsic and extrinsic motivation. The article concludes with a discussion on intervention issues.

1. INTRODUCTION

Motivation represents one of the most important variables in sport. In fact, coaches and athletes agree that motivation is one of the key elements that will facilitate not only performance but also a positive experience in the sport area. But what is motivation? More specifically, what are intrinsic motivation and extrinsic motivation? Why do certain athletes seem more motivated

©2004 Elsevier Inc.
All rights reserved.

than others? Does motivation matter with respect to outcomes? If so, how can coaches and other practitioners best facilitate athletes' motivation? The purpose of this article is to deal with these questions and, thus, to provide a better understanding of the nature and functions of intrinsic and extrinsic motivation in sport.

2. DEFINITIONS AND ASSESSMENT

2.1. On Motivation

Motivation has been the subject of attention from a host of individuals, including playwrights, philosophers, and psychologists. Psychologists, in particular, have looked at the concept for the past two centuries and have provided some general definitions of the construct. A widely accepted definition of motivation is that it represents the hypothetical construct used to describe the internal and/or external forces that lead to the initiation, direction, intensity, and persistence of behavior. Thus, motivation leads to action.

One of the difficulties in defining motivation is that it is not directly observable. Thus, one strategy used to define motivation, implied by the preceding definition, has been to look at behavior to determine whether someone is motivated or not. For instance, a coach would infer that a player is motivated to perfect his or her dribble if the player stays after practices and works on dribble drills.

2.2. On Intrinsic and Extrinsic Motivation and Amotivation

2.2.1. Intrinsic Motivation

Over the years, psychologists have come to realize the existence of different types of motivation. Two broad types that have been studied extensively are intrinsic motivation and extrinsic motivation. Intrinsic motivation refers to engaging in an activity for itself and for the pleasure and satisfaction derived from participation. An example of intrinsic motivation is an athlete who plays basketball because he or she finds it interesting and satisfying to learn new moves with the ball. In 1992, Vallerand and colleagues proposed and showed that there are at least three types of intrinsic motivation: intrinsic motivation to know (engaging in the activity for the pleasure of learning), intrinsic motivation toward accomplishments (engaging in the activity for the pleasure of trying to surpass oneself), and intrinsic motivation to experience stimulation (engaging in the activity out of sensory and aesthetic pleasure).

2.2.2. Extrinsic Motivation

When extrinsically motivated, individuals do not engage in the activity out of pleasure but rather do so to derive some kind of rewards that are external to the activity itself. An athlete who participates in the Olympics to obtain a gold medal and the associated fame and fortune that go with it represents an instance of an extrinsically motivated athlete. Avoiding punishment also pertains to extrinsic motivation.

In 1985, Deci and Ryan published a very influential book titled *Intrinsic Motivation and Self-Determination in Human Behavior*. In their book, they proposed the existence of a number of types of extrinsic motivation that vary in terms of their inherent levels of self-determination. From the lowest to the highest level of self-determination, these types of extrinsic motivation are external regulation, introjected regulation, identified regulation, and integrated regulation. Figure 1 presents the various types of motivation on the self-determination continuum.

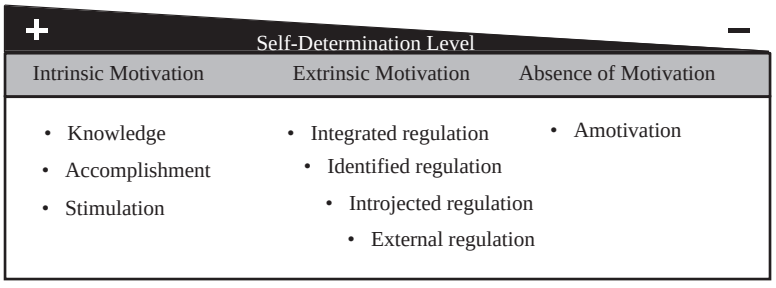


FIGURE 1 The various types of motivation on the self-determination continuum. Adapted from Vallerand and Perreault (1999).

External regulation refers to behavior that is regulated through external means such as obtaining rewards (e.g., medals, trophies) and avoiding constraints (e.g., social pressure). For instance, an athlete might say, "I'm going to today's practice because I don't want the coach to make me sit on the bench during the next game." A second type of extrinsic motivation is introjected regulation. With this type of regulation, the individual has started to internalize the reasons for his or her actions. However, such internalization merely replaces the external source of control with an internal one, that is, self-imposed guilt and anxiety. Thus, an athlete might say, "I go to practices because I would feel guilty if I did not." To the extent that the behavior becomes valued and judged to be important for the individual, and especially to the extent that it is perceived as chosen by oneself, the internalization of extrinsic motives becomes regulated through identification. For instance, an individual might say, "I choose to go to practices because it will help me to perform better during the next game." Finally, an integrated regulation also involves engaging in an activity from an extrinsic perspective in a "choiceful" manner. However, in this case, one's choices are made as a function of coherence with the various aspects of the self. For instance, the integrated athlete might decide to stay home on a Saturday night, rather than go out with friends, so that the athlete will be ready for the next day's soccer game.

2.2.3. Amotivation

Finally, psychologists have also proposed the existence of the construct of amotivation, which refers to a lack of purpose and intentionality in one's action. When amotivated, athletes experience feelings of incompetence and expectancies of uncontrollability and an important loss of motivation. For instance, an amotivated athlete might feel incompetent and start to wonder whether participating in sport still represents a worthwhile option. Amotivation amounts to the relative absence of motivation.

2.3. Assessment of Intrinsic and Extrinsic Motivation and Amotivation

In his 1997 article on the hierarchical model of intrinsic and extrinsic motivation, Vallerand proposed that intrinsic and extrinsic motivation and amotivation, exist at three levels of generality. From the highest to the lowest level in the hierarchy, these are the global, contextual, and situational levels.

Motivation at the situational level refers to a motivational state. It is the motivation that people experience toward a given activity at a specific point in time. For instance, situational motivation toward swimming would be the motivation experienced by a swimmer while training in the pool at 6 o'clock on a Friday morning.

Intrinsic motivation at the situational level has been typically measured in two ways. First, psychologists have used what is called the free choice period. Participants, after having been subjected to experimental conditions, are told that the experiment is over and are provided with the opportunity to engage again in the activity if they so desire. Participants are then left alone and observed unobtrusively. The number of seconds spent on the activity represents the measure of participants' intrinsic motivation toward the experimental task. The more time spent on the task, the higher participants' intrinsic motivation.

This behavioral index of intrinsic motivation is particularly useful in laboratory conditions. However, it is not as useful in real-life conditions, such as sport, where the free choice measure cannot be readily used for practical reasons. Researchers increasingly are turning toward self-report questionnaires to assess participants' motivation. In line with the definition of intrinsic and extrinsic motivation, psychologists have used participants' reasons to participate in the activity as a measure of intrinsic and extrinsic motivation. One self-report instrument that has been used at the situational level with much success in sport is the Situational Motivation Scale (SIMS). Developed by Guay and colleagues in 2000, the SIMS measures four types of motivation: intrinsic motivation without distinguishing the three types, identified regulation, external regulation, and amotivation. Only four subscales of four items each are used to keep the scale brief so as to capture situational motivation in field and laboratory situations. The scale has shown high levels of validity and reliability.

Motivation at the contextual level refers to a relatively stable and general motivational orientation that pertains only to activities that are subsumed under a specific broad life domain such as education, interpersonal relationships, or sport. Thus, individuals come to develop motivational orientations with respect to intrinsic and extrinsic motivation and amotivation in different contexts. A given contextual motivation always refers to one specific life domain. For instance, an

athlete with a high externally regulated motivational orientation toward basketball would generally tend to engage in various basketball situations to receive rewards and/or avoid punishment. Interestingly, a person's contextual motivational orientation may vary from context to context. Indeed, an athlete may have a high level of contextual intrinsic motivation toward his or her sport but a low level of contextual intrinsic motivation toward education.

One instrument that has been used to measure motivation in sport at the contextual level is the Sport Motivation Scale (SMS). In this questionnaire, validated by Pelletier and colleagues in 1995, athletes are asked the following question: "Why do you practice your sport?" Each item represents a perceived reason for engaging in the activity, thereby reflecting the various types of motivation seen previously. Items are assessed using a 7-point Likert-type scale. Sample items appear in Table I. As can be seen, the various types of intrinsic and extrinsic motivation and amotivation are assessed through the reasons for people's participation in sport. There are four items per subscale and seven subscales. The total score of each subscale is a sum of the four items. It should be noted that the concept of integrated regulation is not assessed in the SMS because research has shown that this construct has not yet crystallized in adolescents and young adults. Research using the French and English versions of the SMS reveals that the scale has high levels of validity and reliability. The scale has also been validated in several other languages, including Chinese, Italian, Japanese, Russian, and Spanish.

Finally, motivation at the global level refers to a general motivational orientation to interact with the environment in an intrinsic, extrinsic, or amotivated way. Contrary to contextual motivation, global motivation does not pertain specifically to one given life domain or set of activities but rather is akin to a broad personality trait. Thus, certain individuals may have a global external regulation orientation. Such a personality orientation would predispose individuals to have an external regulation toward several life contexts, including sport. Guay, Pelletier, Vallerand, and others have developed the Global Motivation Scale to assess the seven types of motivational constructs that appear in Fig. 1 (except for integrated regulation) at the global level. It has shown high levels of reliability and validity.

TABLE I
Sample Items from Each of the Seven Subscales of the Sport Motivation Scale

"In general, why do you practice your sport?"	
• Intrinsic motivation—knowledge	"For the pleasure of discovering new training techniques"
• Intrinsic motivation—accomplishments	"For the pleasure that I feel while executing certain difficult movements"
• Intrinsic motivation—stimulation	"For the pleasure I feel in living exciting experiences"
• Identified regulation	"Because it is one of the best ways to maintain good relationships with my friends"
• Introjected regulation	"Because I must do sports to feel good about myself"
• External regulation	"For the prestige of being an athlete"
• Amotivation	"I don't know anymore; I have the impression of being incapable of succeeding in this sport"

3. ON THE DETERMINANTS OF INTRINSIC AND EXTRINSIC MOTIVATION

Because motivation is present in people at three levels of generality, it is important to look at the determinants of motivation at each of the three levels: global, contextual, and situational.

3.1. Where Does Situational Motivation Come From?

Research reveals that at least three types of variables produce an important effect on situational motivation: the task, the person, and the environment.

3.1.1. The Task

Clearly, tasks differ in terms of their intrinsic properties in that certain tasks are more enjoyable than others. For instance, playing basketball would be more interesting to most basketball players than would running up the stairs of a stadium. Thus, it is

not surprising that certain tasks generate higher levels of situational intrinsic motivation than do others.

Yet individual differences and environmental variables may affect perceptions of tasks in such a way that even dull tasks may, at times, be experienced as enjoyable. Thus, much research has focused on the impact of environmental and individual difference variables on intrinsic and extrinsic motivation.

3.1.2. The Environment

Historically, research on intrinsic and extrinsic motivation started during the early 1970s by looking at the effects of rewards on situational intrinsic motivation. In what is recognized as the first contemporary publication on intrinsic motivation, Deci showed in 1971 that receiving rewards to engage in an enjoyable activity produces a decrease in situational intrinsic motivation. Specifically, individuals who received monetary rewards to engage in a fun activity spent less time on the activity subsequently than did individuals who did not receive rewards to engage in the same activity. Subsequent research showed that the negative effects of rewards are especially potent when the rewards are expected and provided to entice participants to engage in the activity. Although there have been some controversies over the years, a meta-analysis of more than 100 studies conducted by Deci and colleagues in 1999 provided overwhelming support for the negative effects of rewards on situational intrinsic motivation. Research in sport, by researchers such as Orlick and Mosher in 1978, has replicated these findings and shown that trophies and awards do undermine intrinsic motivation.

Over the years, research in sport has expanded and looked at the impact of other situational factors on intrinsic motivation. Situational factors are variables that are present at a given point in time. Research has shown that factors such as competition, winning or losing, and performing well or poorly affect situational intrinsic and extrinsic motivation. Specifically, competing to win at all costs, as well as losing or not playing well, produces a decrease in intrinsic motivation and identified regulation and an increase in introjected and external regulation and amotivation. However, noncompetitive situations, as well as winning and doing well, produce the opposite effect.

Why do factors as diverse as rewards, competition, and losing influence athletes' intrinsic and extrinsic motivation? According to Deci and Ryan's self-determination theory, social factors influence motivation through their impact on one's perceptions of competence, autonomy, and relatedness. Because individuals experience the needs to feel competent, autonomous, and related to significant others in their interaction with the environment, activities that allow people to satisfy these needs will be engaged in choicefully and on a regular basis. In other words, athletes experience important and powerful needs to feel connected to others in the sport milieu, to behave in an effective manner in that milieu, and to experience a sense of personal initiative in the process. Variables that nurture those needs help intrinsic motivation and identified regulation to flourish. Variables that thwart those needs produce negative effects on the latter types of motivation and facilitate the development of external and introjected regulation and amotivation. For instance, performing badly in a sport leads one to feel incompetent, and this in turn fuels a loss of intrinsic motivation and identified regulation and a concomitant increase in external and introjected regulation and amotivation. Research using sophisticated statistical analyses, such as structural equation modeling and path analysis, has shown that the impact of the environment on intrinsic and extrinsic motivation is indeed mediated by perceptions of competence, autonomy, and relatedness.

3.1.3. Individual Differences

We have seen how social factors can have a very important impact on athletes' situational motivation. Yet athletes who play on the same team, and who are subjected to the same environmental conditions, often display different levels of situational intrinsic and extrinsic motivation. Why is this the case? Psychologists propose that these differences take place because athletes may differ with respect to their motivational orientations. A motivational orientation typically refers to a more or less stable predisposition to be intrinsically or extrinsically motivated or even amotivated. Such an orientation can be present at the contextual or global level. For instance, an athlete with an intrinsic contextual motivation toward basketball is more predisposed to display higher levels of intrinsic motivation in various basketball situations than is an athlete with a contextual motivation characterized by low intrinsic motivation. This refers to what Vallerand called the top-down effect. It is important to note that such effects appear to hold only for contextual motivational orientations related to the activity in which the individual is engaging. Thus, having a contextual intrinsic motivational orientation toward hockey would help an athlete to maintain a high level of situational

intrinsic motivation toward hockey. However, having a contextual intrinsic motivational orientation toward school would not produce such an effect. Finally, global motivation will influence situational motivation only when the activity being engaged in is rather ambiguous and unrelated to a given contextual motivation. In such instances, global motivation has room to exert its influence on situational motivation.

3.2. On the Determinants of Contextual Motivation

One's contextual motivation toward a specific set of activities can also be affected by the same three factors just reviewed. Thus, task factors are also important influences on contextual motivation. For example, a baseball player's contextual motivation will likely be more intrinsic for actual baseball games than for weight training.

Social factors at the contextual level have been also found to affect contextual motivation. Such factors are those that are present on a general or recurrent basis in one life context (e.g., the sport in which one participates) but not necessarily in other life contexts (e.g., education, interpersonal relationships). Contextual factors studied in sport can be loosely classified into three categories: structural variables (e.g., the types of leagues in which athletes play), motivational climates that prevail in team settings, and interpersonal variables (e.g., the coach's managerial style). Highly structured and competitive league environments, performance-oriented team climates, and playing for a highly controlling coach who does not leave room for autonomy in the athletes all produce decreases in intrinsic motivation and identified regulation while producing increases in introjected and external regulation and amotivation. However, more relaxed league structures, mastery-oriented team climates, and playing for an autonomy-supportive coach who provides understanding, care, and room for decision making on the athletes' part all produce the opposite effect on contextual motivation.

A third source of influence on contextual motivation originates from individual differences. Just like situational motivation, contextual motivation is also subjected to top-down effects originating from individual differences. However, in this case, the influence comes from global motivation. Thus, athletes who entertain a high level of global intrinsic motivation generally possess high levels of intrinsic motivational orientations in various life contexts, including sport.

A final influence on contextual motivation comes from the cumulative effects of experiencing various motivational states on a regular basis. This refers to the recursive bottom-up effect. For example, a tennis player who interacts daily with an autonomy-supportive coach will likely experience daily doses of situational intrinsic motivation that, over time, should have positive recursive effects on the athlete's contextual motivation toward tennis. In the long run, the athlete's contextual motivation will likely become increasingly intrinsic in nature.

3.3. Motivational Determinants at the Global Level

Very little is known about the determinants of global motivation, especially as it relates to the realm of sport. One source of influence that has been documented deals with the recursive bottom-up effect discussed previously. Thus, experiencing high levels of intrinsic motivation at the contextual level in sport leads to the development of a more intrinsic global motivation. Of course, this effect takes place to the extent that this life domain is important for the individual, as would be the case for athletes.

An additional source of influence comes from the impact of global factors. These refer to social factors that are present in most contexts of a person's life on a relatively permanent basis. Global factors help to shape a person's global motivation. In this regard, the motivational impact on global motivation of enrolling at a young age in soccer boarding schools, such as those found in several countries around the world, deserves attention. Such boarding schools represent a global environment in that athletes not only learn how to play soccer, in the hope of becoming professional soccer players in the future, but also go to school and actually live at the boarding schools, away from their families, for years. Therefore, boarding schools should profoundly affect the development of young individuals' personalities, including their global motivation. The type of effect induced, however, will likely depend on the climate that prevails in such schools. If the climate is one of autonomy support, care, and understanding, athletes will likely develop a self-determined global motivation. However, tightly controlling boarding schools will likely lead to the development of non-self-determined global motivation. Thus, being subjected to such global factors shapes athletes' global motivation and personalities. Such a global motivation,

in turn, greatly affects athletes' motivation toward sport through the top-down effect.

4. ON MOTIVATIONAL OUTCOMES

4.1. Motivation Causes Outcomes

It is intuitively thought that high levels of motivation will likely lead an athlete to play hard, persist more at training, and display more intense behavior compared with a less motivated athlete. But is it the case? Does motivation actually cause the outcome? The answer to this question is an unequivocal yes. Indeed, several studies have shown that experimentally inducing intrinsic motivation, as opposed to extrinsic motivation, leads to several positive outcomes, including enhanced creativity, performance, and persistence. In other words, motivation matters.

4.1.1. Outcomes as a Function of the Various Types of Motivation

The type of motivation that one entertains toward an activity also matters in that outcomes are decreasingly positive as one moves from intrinsic motivation to amotivation (see Fig. 1). Because the various types of motivation are located on a continuum ranging from high to low self-determination, and because self-determination is associated with enhanced psychological functioning, a corresponding pattern of consequences ensues. Thus, the most positive outcomes result from the self-determined forms of motivation (intrinsic motivation and identified regulation), whereas negative outcomes follow from the least self-determined forms of motivation (external and introjected regulation and especially amotivation). This pattern of results has been obtained in sport and physical activity with the following outcomes: concentration, sportsmanship, and positive affective experiences such as flow, interest, enjoyment, satisfaction, and less anxiety as well as behavioral outcomes that include expanded effort, behavioral intentions, and actual levels of persistence, adherence in sport, and performance.

It should also be noted that these findings have been obtained with different populations, including elite and recreational athletes, athletes of various ages, physically challenged athletes, intellectually challenged children, and coaches. Thus, it would appear that the role of motivation in producing outcomes is well established and that the more self-determined the motivation, the more positive the outcomes.

4.1.2. Outcomes as a Function of Levels of Generality

Just like motivation, outcomes are experienced at various levels of generality. For instance, having the sudden intention to drop out of basketball is not the same as making such a decision after several weeks or months of reflection. Similarly, feeling momentarily dejected after a loss in a tennis tournament is different from experiencing a generalized negative mood each time one plays tennis. Why is this the case?

Psychologists have shown that outcomes not only differ in terms of generality but also are produced by motivations that differ in terms of their levels of generality. Thus, situational types of outcomes, such as feeling momentarily disappointed and not wanting to play any more at that point in time, originate from situational motivation. On the other hand, more contextually generalized thoughts, feelings, and behaviors that go beyond the moment and pertain to a specific area, such as tennis or football, are produced by sport contextual motivation specific to tennis or football. Finally, broad-level types of outcomes, such as full-fledged depression and apathy experienced across life domains, are typically the fruit of global motivation.

5. INTERVENTIONS

It is well known that participation in sport and exercise produces several physical and psychological benefits. Thus, it is unfortunate to see that a large number of children drop out of sport at a young age. For instance, in Canada, there is a 30 to 35% annual dropout rate in young competitive swimmers, some of whom have great potential. In 1999, Vallerand and Losier proposed the following causal sequence: Social Factors → Psychological Mediators → Motivation → Consequences. Such a sequence integrates the determinants and outcomes of motivation and, as such, charts the nature of the psychological processes through which athletes come to experience various outcomes, including dropping out of sport (Fig. 2). Indeed, research has shown that coaches who display controlling behavior (a social factor) undermine young athletes' sense of competence, autonomy, and relatedness (the psychological mediators). This, in turn, undermines their self-determined motivation,

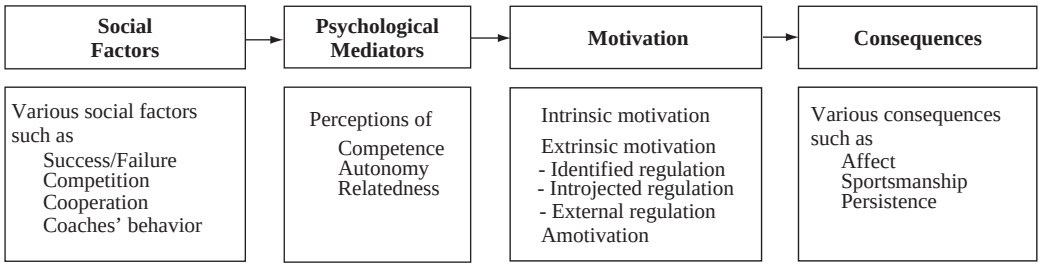


FIGURE 2 The motivational causal sequence: From social factors to outcomes. Adapted from Vallerand and Losier (1999).

leading to their dropping out of sport (a consequence). These findings have been obtained with both male and female athletes who engage in various sports such as handball and swimming.

If motivation has a causal influence on persistence, it might be possible to increase athletes' motivation and, in turn, their persistence toward sport by using the causal sequence depicted previously. This could be done by modifying social factors known to negatively affect motivation such as coaches' controlling interacting style. Although very few intervention studies have been conducted in the realm of intrinsic and extrinsic motivation, one sport study has followed this lead. In 2003, Pelletier and colleagues reported the development of an intervention program dealing with swimmers' motivation and persistence. This 18-month program mainly focused on the following:

- Helping coaches to become more autonomy supportive, thereby enhancing athletes' feelings of competence and autonomy
- Teaching athletes how to deal with the increased levels of competence and autonomy and to become more proactive in their sport environment

It was anticipated that changing coaches' interacting style to a more autonomy-supportive approach would trigger the previously described causal sequence and would lead athletes to experience higher levels of competence and autonomy, higher levels of intrinsic motivation, and higher levels of persistence in swimming.

Results supported this hypothesis. Approximately 1½ years into the intervention program, athletes perceived their coaches as significantly less controlling and as more autonomy supportive. In turn, athletes' levels of perceived competence and intrinsic motivation toward swimming increased, leading to several adaptive consequences. The levels of attendance in the practices increased dramatically, and the annual dropout rate deceased from 35% to less than 5%. A control group of swimmers who were not part of the intervention still displayed a 35% dropout rate. In addition, 85% of athletes who were part of the intervention reached the national performance standards, four were selected to the Canadian Olympic swimming team, and one won a silver medal in the Seoul Olympics. Although preliminary, results from this intervention study underscore the potential for intervention programs that use the proposed motivational sequence as a guiding light.

6. CONCLUSION

Sport motivation is a complex and multidimensional phenomenon. It is influenced by numerous factors and can lead to a host of consequences. Furthermore, not only are intrinsic motivation and identified regulation important for allowing athletes to experience satisfying participation in sport, but these self-determined forms of motivation also lead to higher levels of achievement. In this vein, interventions that promote these forms of motivation have been found to be effective. The road to enjoyable and successful participation in sport has been mapped out.

See Also the Following Articles

Goal Setting and Achievement Motivation in Sport ■ Motivation and Culture

Further Reading

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*, *18*, 105–115.

Deci, E. L. (1975). *Intrinsic motivation*. New York: Plenum.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.

Deci, E. L., & Ryan, R. M. (2002). *Handbook of self-determination research*. Rochester, NY: University of Rochester Press.

Goudas, M., Biddle, S., Fox, K., & Underwood, M. (1995). It ain't what you do, it's the way that you do it! Teaching style affects children's motivation in track and field lessons. *The Sport Psychologist*, *9*, 254–264.

Guay, F., Vallerand, R. J., & Blanchard, C. (2000). On the assessment of the situational intrinsic and extrinsic motivation: The Situational Motivation Scale (SIMS). *Motivation and Emotion*, *24*, 175–213.

Orlick, T. D., & Mosher, R. (1978). Extrinsic awards and participant motivation in a sport related task. *International Journal of Sport Psychology*, *9*, 27–39.

Pelletier, G. L., Fortier, M. S., Vallerand, R. J., & Brière, N. M. (2001). Associations between perceived autonomy support, forms of self-regulation, and persistence: A prospective study. *Motivation and Emotion*, *25*, 279–306.

Pelletier, L. G., Fortier, M. S., Vallerand, R. J., Tuson, K. M., Brière, N. M., & Blais, M. R. (1995). Toward a new measure of intrinsic motivation, extrinsic motivation, and amotivation in sports: The Sport Motivation Scale (SMS). *Journal of Sport & Exercise Psychology*, *17*, 35–53.

Sarrazin, P., Vallerand, R., Guillet, E., Pelletier, L., & Cury, F. (2001). Motivation and dropout in female handballers: A 21-month prospective study. *European Journal of Social Psychology*, *31*, 1–24.

Vallerand, R. J. (1997). Toward a hierarchical model of intrinsic and extrinsic motivation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (pp. 271–360). San Diego: Academic Press.

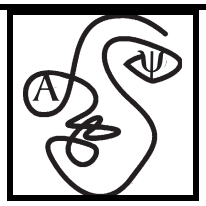
Vallerand, R. J. (2001). A hierarchical model of intrinsic and extrinsic motivation in sport and exercise. In G. Roberts (Ed.), *Advances in motivation in sport and exercise* (2nd ed., pp. 263–319). Champaign, IL: Human Kinetics.

Vallerand, R. J., & Fortier, M. S. (1998). Measures of intrinsic and extrinsic motivation in sport and physical activity: A review and critique. In J. Duda (Ed.), *Advancements in sport and exercise psychology measurement* (pp. 83–100). Morgantown, WV: Fitness Information Technology.

Vallerand, R. J., & Losier, G. F. (1999). An integrative analysis of intrinsic and extrinsic motivation in sport. *Journal of Applied Sport Psychology*, *11*, 142–169.

Vallerand, R. J., & Reid, G. (1984). On the causal effects of perceived competence on intrinsic motivation: A test of cognitive evaluation theory. *Journal of Sport Psychology*, *6*, 94–102.

Vallerand, R. J., & Rousseau, F. L. (2001). Intrinsic and extrinsic motivation in sport and exercise: A review using the hierarchical model of intrinsic and extrinsic motivation. In R. Singer, H. Hausenblas, & C. Janelle (Eds.), *Handbook of sport psychology* (2nd ed., pp. 389–416). New York: John Wiley.



Intrinsic Motivation and Self-Determination

Edward L. Deci

University of Rochester, Rochester, New York, USA

1. Introduction
2. Extrinsic Motivation Versus Intrinsic Motivation
3. The Interaction of Extrinsic Motivation and Intrinsic Motivation
4. Extrinsic Motivation
5. Application of Research on Self-Determination
6. Summary and Conclusions
Further Reading

GLOSSARY

autonomy support A quality of a person's interpersonal environment in which others acknowledge the person's perspective, provide choice, encourage initiative, and minimize the use of pressure and control.

controlling environments Environments in which people in positions of authority, expertise, or power pressure others to think, feel, or behave in specific ways; it is the antithesis of autonomy support.

extrinsic motivation Doing a behavior because it will lead to some separate outcome such as a reward, approval from others, or the avoidance of punishment.

integration The fullest type of internalization in which people take in an extrinsic motivation, identify with it, and integrate it into their sense of self.

interpersonal context The ambience or "climate" of a particular setting that can vary from supporting people's autonomy to controlling their behavior.

intrinsic motivation Doing an optimally challenging activity that is interesting, has an internal perceived locus of causality, and is energized by the basic psychological needs for competence, autonomy, and relatedness; such an activity does not require a contingent outcome that is separate from the activity itself.

introjection A form of internalization in which people take in an external control without ever really accepting it as their own and then use it to pressure and control themselves.

Self-determination in human behavior is based in autonomous motivation, which encompasses both intrinsic motivation and integrated extrinsic motivation. Intrinsic motivation involves doing an activity without the necessity of external prompts or rewards because it is interesting and satisfies the basic psychological needs for competence, autonomy, and relatedness. Extrinsic motivation involves doing an activity because it leads to some separate outcome such as a reward, approval from others, or the avoidance of punishment. Intrinsic motivation is invariantly autonomous, whereas extrinsic motivation varies in its degree of autonomy. Extrinsic motivation becomes autonomous through the processes of internalization and integration. Both intrinsic motivation and integrated extrinsic motivation are facilitated in home, school, work, and clinical settings that are interpersonally supportive, that is, where people's perspectives are acknowledged and where people are

©2004 Elsevier Inc.
All rights reserved.

encouraged to experiment, allowed to try their own solutions to problems, provided with choice, and responded to when they initiate. Under such supportive conditions, individuals are more engaged and persistent, perform more effectively, and display higher levels of psychological health and well-being.

1. INTRODUCTION

The delight of discovery and the joy of leisure pursuits are experiences known to most everyone. They are stimulating, satisfying, and sustaining of the behaviors that promote them. Within the field of motivation, these experiences and the psychological processes that underlie them are encompassed by the concept of intrinsic motivation, which was introduced to the field by Harlow in 1950 and began receiving empirical attention two decades later. The concept has become pivotal in our understanding of human motivation and has provided important insights even about people's motivation for life's mundane activities such as studying uninteresting topics and doing boring work.

Intrinsic motivation is readily observable in young children; indeed, play is the "business of their lives." They happily turn random objects into toys, they role-play others such as their parents, and they invent all kinds of games. Intrinsic motivation is also apparent in older individuals, although perhaps not so ubiquitously. Leisure activities such as playing golf, looking at paintings, and pursuing a bargain on a new dress are often intrinsically motivated (i.e., done for the spontaneous satisfaction provided by the activities themselves), and some people are fortunate enough to be intrinsically motivated for their jobs.

Psychologists began to treat the experience of delight and joy (i.e., of interest and enjoyment) as an important motivator following from the realization that a fuller understanding of human motivation required distinguishing among types of motivation rather than just focusing on the amount of motivation. Stated differently, most theories of motivation or behavioral control, beginning with behavioral theories such as that of Hull and continuing through the cognitive theories of Bandura and Seligman, treated motivation as if it differed only in amount and not in type. People who were more motivated for a particular activity (or class of activities) were expected to do the activity more often and to do it more effectively than were people with less motivation. It was the amount of motivation that was considered critical.

The idea that there are different types of motivation implies that if two people have the same amount of motivation for a particular activity but the types of motivation are different, the quality of their engagement in and persistence at the activity will also be different. For example, there are times when employees may be highly motivated to do an activity at work because they are fearful that they will be punished if they do not do it, and there are other times when employees are equally highly motivated for such an activity because they find it interesting and personally important. Any individuals who can relate to this probably recognizes that these different types of motivation, even if the amounts are the same, can have a major effect on the quality of their engagement with the activities and perhaps even on their feelings about themselves.

2. EXTRINSIC MOTIVATION VERSUS INTRINSIC MOTIVATION

The first important differentiation of motivation into types distinguished between extrinsic motivation and intrinsic motivation. Extrinsic motivation refers to doing a behavior because it will lead to some separate outcome such as a reward, approval from others, or the avoidance of punishment. The concept of extrinsic motivators is linked to behaviorist theories that were prominent in psychology during the 1940s and 1950s, although those theories did not use the term "extrinsic motivation." Behavior modification, for example, was based solely on extrinsic motivators—do a desired behavior and get a reward (which behaviorists termed a reinforcement). Behavior modification was in widespread use in education, therapy, and job training during the 1970s, and then its use gradually tapered off.

2.1. Extrinsic Motivation and Behaviorism

The two primary strands of behaviorism, Hullian drive theory and Skinnerian operant theory, both were built around the concept of reinforcement. In both theories, people learn behaviors when they are reinforced. However, the two theoretical approaches define the concept of reinforcement very differently.

In drive theory, reinforcers are events that reduce basic physiological drives such as hunger and thirst. For example, food reinforces behaviors that lead to it while the hunger drive is operative. In contrast, within

operant theory, reinforcers are defined as any separable consequence (e.g., a food pellet, a dollar bill) that increases the likelihood of a particular response being emitted. In operant theory, reinforcers are not tied to underlying drives.

In operant theory, anything is a reinforcer if it increases the rate of a particular behavior, whereas in drive theory, only those things that reduce drives are called primary reinforcers. However, drive theory includes a concept of secondary reinforcers, which are other things (e.g., monetary rewards, smiles) that take on reinforcing value by being paired with primary reinforcers. For example, money can reinforce because it becomes associated with eating and/or drinking.

2.2. Intrinsic Motivation

Intrinsic motivation is based in people's inherent tendency to be proactive, to interact with the world in an attempt to have an effect, and to feel a sense of accomplishment. Indeed, when people are at their healthiest, they are curious, eager to take on challenges, partial to novelty, engaged with interesting tasks or stimuli, and ready to learn. All of these are manifestations of intrinsic motivation, and they are critical for optimal development across the life span.

The concept of intrinsic motivation arose, in part, in response to problems that became apparent within each strand of behaviorism. Accordingly, there were initially two different characterizations of the concept. The first suggested that not all behaviors are derived from the basic drives such as thirst, hunger, and sex. For example, children often forgo eating, even if they have not eaten for a long time, so as to continue playing a game they find to be fun. Accordingly, White asserted that, just as people have biological needs—the so-called drives—that must be satisfied for them to remain healthy, people also have psychological needs that must also be satisfied for them to thrive. White specified the need for competence as being important. Subsequently, Deci and Ryan maintained that there are three fundamental psychological needs—the needs for competence, autonomy, and relatedness—that are operative for all humans and must be satisfied for them to develop and function optimally.

The other approach to defining intrinsic motivation suggested that intrinsically motivated behaviors are those for which the behaviors themselves are rewarding and, thus, do not require separable consequences. This, of course, stands in contrast to the Skinnerian perspective that suggests that all behaviors are controlled by reinforcing consequences that are separate from the behaviors themselves.

When the two characterizations of intrinsic motivation are considered together, the conclusion is that not all behaviors are controlled by reinforcements and that inherent psychological needs for competence, autonomy, and relatedness underlie those behaviors. Other discussions of intrinsic motivation have helped to flesh out the definition. DeCharms used attribution theory to suggest that extrinsic motivation is operative when people perceive the locus of causality for their behavior to be external to themselves, whereas intrinsic motivation is operative when people perceive the locus of causality to be internal or within themselves. For example, when people are doing an activity to get a reward, they would likely see the reward as the cause of their behavior; thus, the perceived locus of causality (PLOC) would be external. When people are doing an activity because they find it interesting and fun, they would likely see the enjoyment as the cause of the behavior, and so the PLOC would be internal. The idea of PLOC relates to the need for autonomy in that people will be satisfying their need for autonomy to the extent that the PLOC for a behavior is internal.

Another characterization that has been widely used in discussions of intrinsic motivation is that being intrinsically motivated for an activity involves doing an activity because people find the activity interesting. The idea is that people do not need a reinforcement or reward if the activity itself interests them. Thus, the concept of "interest" has been considered central to intrinsic motivation. Using the concept of interest highlights that intrinsic motivation exists in the relation between individuals and activities. That is, a person is intrinsically motivated for some activities and not for others, and not everyone is intrinsically motivated for any particular activity. Rather, an individual is intrinsically motivated for a task to the degree that he or she finds the task interesting, and people differ in the extent to which they find any particular task interesting.

A final important piece of the overall characterization of intrinsic motivation states that optimal challenge is central to intrinsic motivation. That is, people are intrinsically motivated to take on optimal challenges because those are the activities that are most interesting, facilitate competence, and promote development.

In sum, intrinsic motivation involves doing an optimally challenging activity that is interesting, has an internal PLOC, and is energized by the basic psychological needs for competence, autonomy, and

relatedness. Such an activity does not require a contingent outcome that is separate from the activity itself.

3. THE INTERACTION OF EXTRINSIC MOTIVATION AND INTRINSIC MOTIVATION

When extrinsic motivation and intrinsic motivation were identified and defined, researchers began to wonder whether the two might interact. Are these two types of motivation additive, such that the amount of one's intrinsic motivation and the amount of one's extrinsic motivation for a particular activity can be summed to yield one's total motivation for the activity? Stated differently, would anything happen to people's intrinsic motivation for an activity if people began to receive extrinsic inducements for doing it? Would intrinsic motivation itself be enhanced, unchanged, or diminished? If the answer is anything other than "unchanged," this implies that the different types of motivation are not simply additive and, thus, that environmental conditions, such as reward contingencies and deadlines, could affect people's intrinsic motivation when they are operating within those environmental conditions.

3.1. Reward Effects

The first studies of intrinsic motivation in humans, published in 1971 by Deci, examined the effects of monetary rewards on the intrinsic motivation of college students. In subsequent studies that included participants from preschool and high school, the effects of other tangible rewards, such as prizes and good player certificates, were also examined. The studies consistently revealed that tangible extrinsic rewards, whether payments or prizes, undermined intrinsic motivation. Specifically, the undermining occurred when the activity was initially interesting, the rewards were salient, and the rewards were contingent on and expected while doing the activity. Deci interpreted these results as indicating that the extrinsic rewards tend to leave people feeling controlled by the rewards (rather than feeling autonomous). Accordingly, their PLOC tends to become more external.

The general finding that tangible rewards tend to undermine intrinsic motivation has been replicated numerous times, but the finding has remained controversial for three decades. Because rewards are the easiest motivators to use, some people resist recognizing their potentially detrimental effects. Nonetheless, in 1999, Deci and colleagues published a meta-analysis involving 128 experiments on reward effects and showed definitively that tangible extrinsic rewards do reliably and clearly undermine intrinsic motivation for the rewarded activity.

3.2. Other External Events

Additional research has shown that other external events, such as threats of punishment, deadlines, directives, and competition, also undermine intrinsic motivation, presumably because people experience them as controls. On the other hand, the opportunity to make choices about what to do and how to do it was found to enhance intrinsic motivation by facilitating a greater experience of autonomy and a more internal PLOC.

The effects of all these studies were interpreted in accord with the postulate that people have a basic need for autonomy. Accordingly, factors that support people's autonomy and experience of an internal PLOC (e.g., choice, acknowledgment) tend to enhance intrinsic motivation, whereas those that leave people feeling controlled with an external PLOC (e.g., tangible rewards, deadlines, pressures) tend to diminish intrinsic motivation.

3.3. Performance Feedback

Other research relates to the basic psychological need for competence. Studies have shown that positive performance feedback (which is sometimes referred to as verbal rewards) enhances intrinsic motivation, that negative feedback diminishes it, and that perceived competence mediates these effects. There is also evidence, however, that the enhancement of intrinsic motivation following praise is less likely to occur for children than for college students. Furthermore, studies have found that increases in perceived competence must be accompanied by feelings of autonomy for the perceived competence to enhance intrinsic motivation. If people feel forced to perform well, their good performance and the positive feedback that might follow it are unlikely to enhance their intrinsic motivation.

3.4. Interpersonal Contexts

A final set of studies has shown that the ambience or "climate" of a particular setting (i.e., the general feeling of the setting) has an effect on people's intrinsic motivation. This factor is frequently referred to as the

interpersonal context, and the critical issue is whether the interpersonal context tends to support people's autonomy or to control their behavior. For example, the coaching climate of some sporting teams tends to encourage athletes to take initiative and experiment, whereas the climate surrounding other teams conveys that athletes must do what the coach tells them to do. Studies show that if the interpersonal context tends to be autonomy supportive, people within it will be more intrinsically motivated, whereas if it tends to be controlling, people will be less intrinsically motivated. For example, if teachers use a lot of controlling language—demands and imperatives as well as words such as "should," "must," and "have to"—their students tend to be less intrinsically motivated. In fact, this issue has been studied in the laboratory, and it seems that if environmental factors such as rewards are administered in relatively autonomy-supportive ways, they are less likely to undermine intrinsic motivation. Furthermore, if factors such as positive feedback are administered in relatively controlling ways, they are less likely to enhance intrinsic motivation.

3.5. Summary

Research has shown that intrinsic motivation and extrinsic motivation are not additive; instead, they are interactive. That is, the addition of extrinsic motivators may either diminish or enhance intrinsic motivation. As such, the social environments within which people operate will significantly affect the degree to which they are intrinsically motivated for the target activity. Contextual factors that allowed satisfaction of the basic psychological needs for competence and autonomy were found to enhance intrinsic motivation, whereas those that thwarted satisfaction of these basic needs were found to undermine intrinsic motivation (Table I).

TABLE I
Summary of the Effects of External Motivators on People's Intrinsic Motivation

Monetary rewards	Tend to decrease intrinsic motivation
Other tangible rewards	Tend to decrease intrinsic motivation
Deadlines	Tend to decrease intrinsic motivation
Threats of punishment	Tend to decrease intrinsic motivation
Competition against others	Tends to decrease intrinsic motivation
Directives and evaluations	Tend to decrease intrinsic motivation
Negative feedback	Tends to decrease intrinsic motivation
Controlling interpersonal climates	Tend to decrease intrinsic motivation
Choice	Tends to increase intrinsic motivation
Acknowledgment	Tends to increase intrinsic motivation
Encouragement for self-initiation	Tends to increase intrinsic motivation
Positive feedback	Tends to increase intrinsic motivation
Autonomy-supportive interpersonal climates	Tend to increase intrinsic motivation

3.6. Intrinsic Motivation and Outcomes

Numerous studies have shown that intrinsic motivation was associated with better conceptual learning, greater creativity, more cognitive flexibility, greater behavioral persistence, and more enhanced well-being than was extrinsic motivation. For example, experiments showed that autonomy-supportive environments led both elementary school and college students to learn better than did controlling environments, especially when the material was complex and required conceptual understanding.

Other research conducted in real-world settings such as homes, sport teams, and work groups has revealed that an autonomy-supportive environment catalyzes not only intrinsic motivation but also performance, level of satisfaction, and general well-being. Indeed, many studies have shown strong positive relations between being intrinsically motivated and having a higher level of psychological health and well-being. Such research makes clear why the differentiation of types of motivation is so important and confirms that the distinction between intrinsic motivation and extrinsic motivation is an empirically useful one.

It is critical to remember that intrinsic motivation is operative only for activities that hold intrinsic interest for individuals, for example, those that are novel and challenging. For activities that are not interesting, intrinsic motivation does not represent a potent motivator. Still, there are many activities that parents, teachers,

managers, coaches, and physicians believe are important for people, even if the activities are not interesting. To understand the motivation for these activities, one must examine the dynamics of extrinsic motivation.

4. EXTRINSIC MOTIVATION

As it turns out, many of the activities people do are not intrinsically motivated, especially from the time they move out of early childhood and face increasing demands to assume social roles and accept responsibilities. Children must begin to interact congenially with playmates, then to do schoolwork that they do not find interesting, and eventually (as adults) to hold gainful employment and function within the laws of society. Evidence shows, for example, that even during children's years in elementary school, their intrinsic motivation tends to become weaker with each passing year.

Recall that extrinsic motivation is operative when people do an activity to attain some separable outcome such as a pat on the back or a monetary bonus. When children are not intrinsically motivated for an important activity, parents or teachers must rely on the inducement of extrinsic motivators. Although extrinsic motivators have typically been found to be controlling rather than supportive of autonomy, there is still the important question of whether it is possible for people to be autonomous or self-determined when doing a task that is uninteresting and, therefore, was extrinsically prompted. The reason this question is so critical is that if it were possible for people to be autonomous while doing an activity for extrinsic or instrumental reasons, perhaps some of the performance and adjustment advantages that have been found for intrinsic motivation might also accrue when people are autonomously extrinsically motivated.

Deci and Ryan's self-determination theory (SDT) proposed that intrinsic motivation is invariantly self-determination but that extrinsic motivation can vary greatly in the degree to which it is self-determination. For example, a boy who takes out the garbage only because he knows that his mother will praise him for doing so is extrinsically motivated because the behavior is instrumental to the separate consequence. A girl who takes out the garbage because she personally values family harmony and believes that her participation in doing this task will help to maintain that harmony is also extrinsically motivated. The second instance, although still extrinsic, is characterized by a greater sense of personal endorsement and choice than is the first instance, which involves compliance with the control that has been explicitly or implicitly created by the boy's mother. These two cases of extrinsically motivated behavior vary in the degree to which they are relatively autonomous.

4.1. Internalization of Extrinsic Motivation

Within SDT, extrinsically motivated behaviors can become more autonomous through a process in which people internalize the values and regulations associated with the behaviors. Internalization is the process of taking in a value or regulation and making it one's own. When the process functions optimally, individuals will have transformed an externally regulated extrinsic motivation into an internally regulated extrinsic motivation by integrating the regulation and its value into their sense of self.

Internalization has long been an important concept in many psychological theories. The idea is that the challenge of becoming self-regulating of activities that are important for people but are not themselves personally compelling can be met through the process of internalization. To be self-regulating, people must make internal what was initially external. The thing that makes SDT's conceptualization of internalization different from most others is that it includes different types or degrees of internalization. In other words, SDT proposes that regulations can be internalized more or less fully, such that people will, to differing degrees, accept the regulations as personally important for themselves and, thus, be more or less autonomous in enacting them. As such, there will be differences in the extent to which people do a behavior because they want to rather than because they believe that they have to, and this will reflect the degree to which the regulation has been internalized.

Thus, SDT views internalization in terms of a continuum that describes how fully the person has transformed an external prompt into an internal regulation that will allow volitional or "choiceful" behavior. This continuum, therefore, illustrates how instrumental (i.e., extrinsic) motivation for a behavior can range from passive compliance to active personal commitment.

One reason why this relative autonomy continuum is so important is that it addresses an important problem about internal regulation of behavior. Specifically, there are behaviors that people force themselves to do because they think they should do them and know they

will feel guilty if they do not. The regulation of these behaviors is certainly internal to these people, but it does not exhibit the qualities of volition or autonomy that, for example, are so evident in intrinsic motivation. SDT uses the concept of introjection to refer to the type of internalization that leads to this internally pressured regulation. In contrast, there are behaviors that people do not find interesting but that have become meaningful and important for their own self-selected goals and personal life plans. As such, people do these behaviors quite autonomously, even though the behaviors themselves are not inherently satisfying. In SDT, this is said to occur as people identify with the importance of the activities for themselves and then integrate that identification with other aspects of themselves. When this has occurred, people will feel fully autonomous as they enact an extrinsically motivated behavior because it would then emanate from their integrated sense of self.

The former example, of introjection, is about controlling oneself, and the process bears considerable similarity to being controlled by other people. The latter example, in contrast, is more about making a choice to do the activity because, all things considered, doing so feels desirable and right. Thus, in the latter case, people experience a truly internal PLOC, even though the behavior is still instrumental, that is, done for reasons other than the enjoyment of the activity itself. This important contrast, between what might be thought of as self-control and self-regulation, is highlighted in the SDT model of extrinsic motivation.

More specifically, the theory states that there are various forms of extrinsic motivation: one where the regulation is external to the person, one where the regulation has been merely introjected and so must be buttressed by internal reward or punishment (e.g., guilt) contingencies, one where the person has identified with the importance of the behavior for himself or herself, and one where the regulation has been fully integrated with other aspects of the person's self. Although the latter three of these all represent internal motivation, they vary in the degree to which the behaviors they motivate are autonomous. Integrated regulation represents a relatively full sense of volition and personal commitment, with introjection representing a relative lack of these qualities (Fig. 1).

Intrinsic motivation and fully integrated extrinsic motivation are the two bases for autonomous or self-determined behaviors. More than three decades of research have now shown that the quality of people's experience and performance vary as a function of the degree to which a behavior is autonomous or self-determined. Thus, when people have identified with and integrated the regulation of an extrinsically motivated behavior, the behavior shares many of the qualities of behaviors that are intrinsically motivated. Not only do people feel a sense of choice and experience an internal PLOC, but more importantly, autonomous extrinsic motivation, like intrinsic motivation, is positively related to psychological well-being as well as to learning outcomes and effective performance, especially on activities that require a deeper or fuller engagement with the activity.

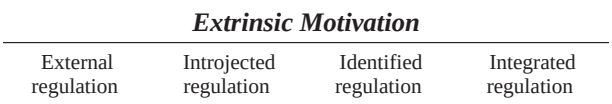


FIGURE 1 Types of extrinsic motivation based on the degree to which a regulation and its underlying value have been internalized. They range from external regulation (least autonomous) to integrated regulation (most autonomous).

4.2. Internalization in Social Contexts

Various studies have examined how the social conditions within which a regulation is internalized affect how fully it is internalized and, thus, how autonomous the subsequent behavior will be. One factor that affects how self-determined people will become for behaviors that were initially externally prompted is whether the behaviors were clearly valued by others to whom they feel connected. This suggests, for example, that the starting point for facilitating internalization in children, students, or employees is providing them with a sense of personal relatedness, that is, the feeling of being cared about. In other words, it is the basic psychological need for relatedness that leads people to be willing to take on the regulation of behaviors that are valued by significant others or relevant reference groups.

In addition, people must feel sufficiently competent at the focal behavior to do it satisfactorily; that is, people will be more likely to internalize a goal if they have the relevant skills and understanding to become successful at it. Thus, supporting competence by offering optimal challenges and providing effectance-relevant feedback will also facilitate internalization. Finally, for people to become autonomous with respect to a behavior, they must grasp its personal meaning for themselves; that is, they must have their own perspective acknowledged and

must feel a sense of choice about doing the behavior. This suggests that whereas people may at least partially internalize a behavioral regulation if they feel a sense of (or desire for) relatedness to a relevant other or group and if they feel effective in doing the behavior, they are unlikely to fully integrate the regulation unless the relevant others in their social world provide autonomy support with respect to that behavior. Notably, then, whereas supports for autonomy and competence are particularly important for maintaining intrinsic motivation for an activity, supports for relatedness are also important for internalization. In short, when people feel supports for relatedness, competence, and autonomy with respect to an extrinsically motivated behavior, they are likely to internalize its value and regulation quite fully.

In sum, self-determination, which is based in intrinsic motivation and integrated extrinsic motivation, has been associated with a variety of positive performance and adjustment outcomes and has been found to depend on interpersonal supports for relatedness, competence, and autonomy. In these regards, then, one can now see how the study of joy and delight provided a starting point for an examination of how to facilitate optimal motivation for even the mundane tasks that are encountered each day in people's lives. By providing the kinds of supports that help to maintain intrinsic motivation, one can also facilitate internalization and integration that will enhance the quality of motivation and engagement with the more mundane activities as well as the more interesting ones.

5. APPLICATION OF RESEARCH ON SELF-DETERMINATION

In various domains of life, there are people who have responsibility for motivating others—their children, their students, subordinates, patients, athletes, and the like. This can be a daunting challenge. Because research has shown that autonomous motivation, based in intrinsic motivation and integrated extrinsic motivation, is unequivocally linked to superior conceptual learning, greater creativity, more flexible problem solving, and better psychological health and well-being, it seems clear that working to facilitate autonomous motivation, rather than taking the easier route of relying on controls, is indeed important.

Numerous studies have examined the effects of supporting autonomy versus controlling behavior in a variety of real-world domains. We now turn to a brief consideration of a few of these studies in some of the more important domains of life.

5.1. Autonomy Support in Parenting

Supporting autonomous motivation and promoting healthy development begin in the home. Parents have the task of providing a context that will allow their children to experience a sense of autonomy, competence, and relatedness, and it is those contexts that foster autonomous motivation, internalization of values, and psychological well-being.

Parents, of course, are genuinely concerned about their children acquiring the skills and values that are necessary for functioning effectively in school, on the playground, and in their subsequent careers and personal development. Parents are also hopeful that their children will behave appropriately in a variety of everyday settings so as not to bring rejection or embarrassment on themselves and their families. At times, parents' concerns and desires can prompt conflict because their expectations can conflict with the interests and desires of the children. Children are not always eager to do their homework, take out the garbage, or attend the religious services in which their parents are invested. This resistance may prompt an urge in parents to apply pressure in an attempt to ensure that their children do the behaviors that the parents believe are in the children's best interests. Research makes clear, however, that the children pay a cost in terms of their motivation and self-reliance to the extent that parents' motivational strategies involve pressure and control. Whether the controls are blatant or subtle, and whether they are coercive or seductive, they seem to have some negative psychological consequences for children.

One common approach to motivation and discipline used by many parents is conditional regard, that is, providing more attention and affection when children do what the parents value and withdrawing affection and regard when the children do not. This parenting strategy turns out to be quite controlling, however, and research has shown that it leads children to merely introject behavioral regulations rather than integrate them. In other words, children will tend to do the behaviors on which affection and regard were made contingent, but they will experience a sense of inner compulsion and anxiety about the behaviors rather than a sense of choice and desire. Furthermore, this pressured form of regulation is associated with poorer well-being—with lower and more fragile self-esteem,

greater feelings of guilt and shame, and the like. In addition, to the extent that parents use conditional regard as a socializing strategy, their children still harbor resentment toward the parents many years later. Other more overtly coercive forms of control have been found to have even more negative effects.

The alternative to pressure and control, of course, is autonomy support, which has been found to enhance motivation, performance, and well-being. In one interview study, for example, Grolnick and Ryan went into the homes of late elementary school-aged children. The researchers conducted separate interviews with mothers and fathers of the children regarding how the parents motivate the children to do schoolwork and to carry out chores around the house. Results of the study made clear that when parents were autonomy supportive—when they took their children's perspective into consideration, encouraged initiative, and provided choice—the children were more autonomously motivated, felt better about themselves, and were rated by teachers as being more engaged in their schoolwork and more competent.

Autonomy support represents a challenge for parents. To try to understand children's behavior from their perspectives can be a formidable undertaking when parents have their own perspectives and desires. To allow and support children's choices when it is easier to impose parents' own choices is asking a lot of parents, especially if they are stressed or tired from dealing with demands in other domains of their lives. Yet the evidence is clear that it is only through autonomy support that children will come to the point of accepting responsibility for the behaviors that are in their best interests. In other words, pressure and control can prompt compliance in the immediate situation (although it sometimes prompts rebellion), but even when it prompts compliance, it is a relatively ineffective socializing strategy because the behavior tends not to persist over the long term. Autonomy support, in contrast, promotes internalization and integration of the behaviors so that they become part of the children's own values and sense of self and, thus, are more likely to be maintained over time.

It is very important to recognize that autonomy support is not the same thing as permissiveness. To be autonomy supportive does not mean to allow children to do anything they want. Rather, autonomy support involves the use of limits and discussions about the reasons why certain behaviors are important. The important thing, however, is the style in which that is done. When parents demand that their children accept limits and punish them for not accepting these limits, the parents are creating a power struggle that is doomed to fail. For limits to work effectively, it is important to acknowledge the children's point of view, for example, recognizing that they might not want to behave in the way that is being discussed. It is also important to provide a meaningful rationale for doing the behaviors and to refrain from using interpersonal pressure. It is the pressure, rather than the limits, that creates the problem. As such, autonomy support means being neither pressuring nor permissive; instead, it involves working with children to support their acceptance of the values and behaviors that are important for their own development and well-being.

5.2. Autonomy Support in Education

Several studies, using questionnaires, observational methods, and experimental designs, have examined the facilitation of motivation and performance in students. Studies have shown that there is substantial variability in public school classrooms regarding the degree to which teachers tend to support students' autonomy versus control their behavior. Furthermore, evidence from research in elementary school classrooms indicates that when teachers are more supportive of students' autonomy—when teachers provide meaningful explanations for why certain behaviors are important, allow students to experiment with alternative ways of doing things, and encourage students to take initiative and try solving their own problems—students display higher levels of intrinsic motivation, curiosity, and desire for challenge and also are more likely to internalize extrinsic motivation in a relatively full way. One study showed that within 2 months in classrooms, students had made an accommodation to the teachers' styles such that students of relatively controlling teachers had become less intrinsically motivated. Because controlling methods have been related to poorer conceptual learning, poorer performance on problem-solving tasks, less creativity, and lower feelings of self-worth, the evidence is compelling that more autonomy-supportive approaches in classrooms have substantial advantages for learning, performance, and development.

During recent years, there have been some efforts to include the concept of supporting students' and teachers' autonomy within school reform approaches, and initial evidence suggests that this is an important element of effective reform. For example, the Institute

for Research and Reform in Education has carried out districtwide reform programs that have the concept of autonomy support as a central element, and these efforts have resulted in a variety of positive educational outcomes, including decreased dropout rates, enhanced engagement in classroom activities and homework, and improved performance on standardized tests.

At the college and medical school educational levels, comparable results have been reported. For example, in a study of undergraduates learning organic chemistry, results indicated that when instructors were perceived as more autonomy supportive, students became more autonomously motivated over the semester, resulting in better performance on a standardized course exam. Studies of medical students similarly showed that instructors who were more autonomy supportive had students who were more likely to maintain interest, internalize important values associated with the course, feel more competent, and perform better in the practical application of course material.

5.3. Autonomy Support in the Workplace

Work groups, like classrooms, often vary in their climate or ambience. Some managers tend to encourage subordinates to participate in goal setting, performance evaluation, and other important decisions that are relevant to the subordinates' work lives, whereas other managers prefer to make the decisions by themselves and then tell the subordinates what needs to be done. The former approach tends to integrate the planning and doing of the work, with subordinates being involved in both, whereas the second tends to separate them, with the managers planning and the subordinates doing.

Research has shown that when managers have a more autonomy-supportive orientation toward their employees, the employees are more motivated, satisfied, and trusting and also perform more effectively than when managers have a more controlling orientation. In one study, researchers developed an intervention in which a change agent worked with managers in a major corporation to train the managers how to be more autonomy supportive in their interactions with subordinates. Results of the study indicated that over the period of the study, managers did become more autonomy supportive. More importantly, however, the subordinates of those managers who had been involved in the training became more motivated and more satisfied with their work than did comparable employees who worked for managers who had not received the training. It appears that when managers become more autonomy supportive, the change in their orientation and behavior radiates to their subordinates, resulting in improved engagement and effectiveness of the organizational unit.

In other research, Deci and colleagues studied managers and their subordinates in Bulgaria and the United States to confirm whether the positive effects of autonomy support would appear in a collectivist totalitarian society (Bulgaria) as well as in an individualistic democratic society (United States). Results showed that when managers were autonomy supportive, their subordinates experienced greater satisfaction of the basic needs for competence, autonomy, and relatedness, leading to greater engagement in work and better psychological health. Thus, autonomy support seems to be important in the workplace regardless of cultural values or the economic and political systems that are operative in the culture.

5.4. Autonomy Support in Health Care

The importance of supporting autonomy versus controlling behavior for the promotion of maintained health behavior change has also been demonstrated in several studies. Specifically, research has shown that health care practitioners who are more autonomy supportive are more successful in promoting healthy behavior among their patients. Because unhealthy behaviors—tobacco use, poor diets, lack of exercise, and the like—are responsible for well over half a million American deaths each year, facilitating change toward healthier behaving is clearly a critical issue.

Williams and colleagues have examined practitioner autonomy support with respect to a variety of health behaviors. For example, provider autonomy support has consistently been found to promote more autonomous motivation for healthier living and greater feelings of competence about making successful maintained change. In one study, these enhanced experiences of autonomy and competence among patients with diabetes resulted in physiologically indexed improvement in glucose control over a year-long treatment period. In another study, the enhancement of patients' autonomous motivation by autonomy-supportive physicians led patients to be more adherent to long-term

medication prescriptions. Because only approximately half of all prescriptions written by physicians in the United States are followed to a degree necessary for the medications to be effective, promotion of adherence is an important issue for health care. It seems that when providers engage patients as partners in the health care process by encouraging them to be more active in their own care, the patients become more autonomously motivated and are more successful in carrying out these difficult health-relevant behaviors.

Another study examined the autonomy support of providers in a weight loss program where morbidly obese individuals were on a medically supervised, very low-calorie diet. Results indicated that the degree of provider autonomy support made a significant difference. Specifically, patients who experienced the staff as more autonomy supportive reported more autonomous motivation for participating in the program, attended weekly program meetings more regularly, and lost more weight during the program. Furthermore, at a 2-year follow-up, they were exercising more regularly and were more effective in maintaining the weight loss.

In one study, autonomy-supportive principles were used as the central element in developing a medically based counseling intervention for smokers. Using more than 1000 smokers whose educations and incomes were significantly below those of the urban county where they lived, the study evaluated the new counseling intervention relative to the usual care for treating tobacco dependence that is available in the community. Results indicated that the new intervention was substantially more effective than typical community care and that it also linked provider autonomy support directly to patients' autonomous motivation for smoking cessation. This, in turn, was a significant predictor of the patients' 6- and 18-month biochemically validated smoking cessation.

5.5. Autonomy Support in Athletics

Field studies have also shown the importance of supporting autonomy and competence for enhancing autonomous motivation in the domain of sport and exercise. There is a popular image of coaches as tyrants who care only about winning and who pressure their players to work hard and win. Although that might not be the norm, there does appear to be a tendency toward pressure and control in sports. Many parents put strong pressure on their young children to win at Little League baseball, for example, and college athletics has become something of a professional endeavor, at least in some circles. Professional athletics, of course, is increasingly tied into huge salaries, endorsement contracts, and intense pressure to win. Studies of competition and motivation in athletics indicate that this may be counterproductive, however, both for performance and for the health and well-being of athletes. For example, there is evidence that pressuring athletes to win decreases their intrinsic motivation for the sport, and athletic scholarships to college football players have been found to have a negative motivational impact.

An interesting study by Pelletier and colleagues showed that when coaches of competitive swimmers were autonomy supportive, the athletes tended to be autonomously motivated and, in turn, persisted at the sport over a 22-month period. On the other hand, when coaches were controlling, the athletes tended to be either external or introjected in their motivation. Externally motivated swimmers had a high rate of dropout from the sport by the time of a 10-month follow-up, and those who were introjected in their motivation tended to persist through the 10-month follow-up but evidenced a high rate of dropout by the 22-month assessment point. It seems clear, then, that although controlling approaches to motivation can bring about some behavior change, that change is unlikely to persist over the long term, whereas autonomy-supportive approaches are much more likely to promote maintained long-term change.

6. SUMMARY AND CONCLUSIONS

Autonomous motivation, which is the basis for self-determination in human behavior, consists of intrinsic motivation (i.e., doing an activity because it is interesting and does not require a separable reward) and integrated extrinsic motivation (i.e., doing an activity because it has become personally important for the person's self-selected values and goals). Both intrinsic motivation and integrated extrinsic motivation are facilitated in situations that are interpersonally supportive, that is, where parents, teachers, managers, health care providers, and coaches understand and acknowledge the perspectives of their children, students, employees, patients, and athletes, respectively, and encourage them to experiment, allow them to try their own solutions to problems, provide them with choices, and respond to their initiations.

Under conditions that are supportive of individuals' autonomy, those individuals are more engaged with the task and persist longer, perform more effectively

(especially on tasks that require greater creativity, flexibility, and conceptual understanding), and display higher levels of psychological health and well-being.

See Also the Following Articles

Motivation and Culture ■ Work Motivation

Further Reading

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191–215.

DeCharms, R. (1968). *Personal causation: The internal affective determinants of behavior*. New York: Academic Press.

Deci, E. L., Betley, G., Kahle, J., Abrams, L., & Porac, J. (1981). When trying to win: Competition and intrinsic motivation. *Personality and Social Psychology Bulletin*, *7*, 79–83.

Deci, E. L., Connell, J. P., & Ryan, R. M. (1989). Self-determination in a work organization. *Journal of Applied Psychology*, *74*, 580–590.

Deci, E. L., Eghrari, H., Patrick, B. C., & Leone, D. R. (1994). Facilitating internalization: The self-determination theory perspective. *Journal of Personality*, *62*, 119–142.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, *125*, 627–668.

Grolnick, W. S., & Ryan, R. M. (1987). Autonomy in children's learning: An experimental and individual difference investigation. *Journal of Personality and Social Psychology*, *52*, 890–898.

Hull, C. L. (1943). *Principles of behavior: An introduction to behavior theory*. New York: Appleton–Century–Crofts.

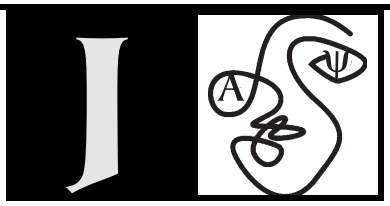
Lepper, M. R., Greene, D., & Nisbett, R. E. (1973). Undermining children's intrinsic interest with extrinsic rewards: A test of the "overjustification" hypothesis. *Journal of Personality and Social Psychology*, *28*, 129–137.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*, 68–78.

Skinner, B. F. (1953). *Science and human behavior*. New York: Macmillan.

White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review*, *66*, 297–333.

Zuckerman, M., Porac, J., Lathin, D., Smith, R., & Deci, E. L. (1978). On the importance of self-determination for intrinsically motivated behavior. *Personality and Social Psychology Bulletin*, *4*, 443–446.



Job Analysis, Design, and Evaluation

Robert Smither

Rollins College, Winter Park, Florida, USA

1. Job Analysis
2. Job Design
3. Job Evaluation
4. Other Issues in Job Analysis, Design, and Evaluation
Further Reading

GLOSSARY

autonomy The level of independence the worker has in deciding how to accomplish the tasks associated with the job.

commensurable factors Behaviors typical of different levels of skill and assigned a number that reflects those differences.

critical incidents Method of job analysis that studies jobs by having knowledgeable individuals provide real workplace examples of successful and unsuccessful job performance.

feedback The degree to which a worker receives information about his or her performance.

growth need strength A worker's need for personal growth through job performance.

hygiene factors Herzberg's term for conditions of work, including salary, management, work environment, and other factors.

job An ongoing set of cognitive or physical activities performed by one person that furthers organizational goals.

job analysis Identification of tasks or behaviors necessary to perform a job successfully.

job design Structuring of tasks and behaviors that creates a job.

job elements A method of analyzing the duties associated with a job.

job enlargement The redesign of a job to make it more interesting by adding duties.

job enrichment The redesigning of a job so that workers have more responsibility for their accomplishment, thereby motivating workers to accomplish their jobs at a higher level of performance.

job evaluation A means of determining the economic value of a job and the setting of salary levels.

job grading A job evaluation method that places jobs into levels or grades.

job ranking A method of job evaluation in which an expert ranks all jobs in an organization or a department in terms of the importance of their contribution to the overall organizational goals.

job simplification The belief that jobs should be designed so that any worker with the requisite skills could perform the job successfully by knowing the procedures to follow.

motivators Tasks and opportunities for advancement and professional growth associated with a job.

physical abilities An analytical approach that focuses on the strength, cognitive abilities, stamina, and other factors required by a specific job.

point method A method of job evaluation in which dimensions of a job are described by subject matter experts and considered the basis for assigning salary levels.

skill variety Abilities that a worker uses to accomplish a job.

subject matter experts (SMEs) People who are familiar with the requirements of a job, such as incumbents, former job performers, supervisors, and job analysts.

©2004 Elsevier Inc.
All rights reserved.

task identity The degree to which the worker sees the outcome of his or her efforts.

task significance The impact of a job on others, both inside and outside the organization.

In its most basic sense, a job is an ongoing set of cognitive or physical activities performed by one person that furthers organizational goals. Information about jobs comes from job analysis, which is the identification of tasks or behaviors, sometimes referred to as the knowledge, skills, and abilities, necessary to perform a job successfully. Job analysis is used for a variety of organizational purposes, including setting selection standards, designing departments and divisions, training, performance appraisal, job design, and job evaluation. Job design refers to the structuring of tasks and behaviors that creates a job and is often undertaken after a job analysis. Job evaluation is the procedure for determining the economic value of a job and the setting of salary levels.

1. JOB ANALYSIS

Researchers often trace the beginnings of job analysis back to the work of Frederick W. Taylor, an industrial engineer at Bethlehem Steel, at the beginning of the 20th century. Taylor's goal was to find what he considered the one best way to perform a specific job and then develop step-by-step procedures for jobholders to follow. Taylor's approach focused solely on the tasks associated with a job—he did not consider social, organizational, or motivational factors in studying jobs. Nonetheless, his studies resulted in significant improvement in efficiency and productivity at Bethlehem Steel.

Taylor's research was guided by the concept of job simplification, or the belief that jobs should be designed to minimize individual differences in workers so that any worker with the requisite skills could perform the job successfully by knowing the procedures to follow.

A few decades after Taylor's original studies, however, the context of most jobs had changed dramatically and job analysis became a more complex procedure than simply identifying tasks and the best way to perform them. For example, Taylor had assumed that worker motivation was primarily economically based. Later research, such as the Hawthorne bank wiring studies or studies regarding the longwall method of coal getting, suggested that factors other than economics affect worker performance.

This increased knowledge about the complexities of worker motivation affected the field of job analysis by broadening the area of study to motivational, social, and organizational considerations in addition to the tasks associated with performing the job. Recently, changing demographics of the workforce, and concerns about making jobs equally available to all persons who can perform them, further broadened the field of job analysis. Although the most basic model of job analysis still consists of identifying tasks, worker characteristics, and working conditions, this identification is now done in the context of a much broader range than the issues Taylor originally considered.

Information about the tasks or behaviors associated with a particular job can be collected in a variety of ways. Typical methods of collecting data include observation of the person performing the job; interviews with the job incumbent, the incumbent's peers, immediate supervisor, or other supervisory personnel with the organization; diaries and other forms of record keeping; and questionnaires. Sources for job analysis information can include, in addition to job incumbents and their supervisors, vendors, customers, or other organizational members who interact with the incumbent.

Throughout the years, researchers have developed a number of approaches for analyzing the duties associated with a job. These procedures, which focus on different aspects of the job, can be quite precise and detailed, and they can focus mainly on either tasks to be performed or characteristics needed to perform the tasks. The job elements approach, for example, focuses on the knowledge, skills, and abilities (KSAs) necessary to perform a job successfully. Information about the KSAs comes from job incumbents and supervisors, who are considered subject matter experts (SMEs). Other approaches take a more generic approach to analyzing jobs. The Position Analysis Questionnaire method, for example, relies on the assumption that there is an underlying taxonomy to all jobs, and that all tasks associated with a job can be classified into six categories.

Another approach to job analysis is to focus on the physical abilities associated with successful job performance. This approach considers strength, cognitive abilities, stamina, and other factors required by a specific job. Physical abilities job analysis has been particularly important in addressing gender issues in personnel recruitment and selection.

Finally, the critical incidents method of job analysis studies jobs by having incumbents and other knowledgeable individuals provide real workplace examples of successful and unsuccessful job performance. From

these examples, job analysts identify the abilities and behaviors necessary to perform a job successfully.

Although each of these methods takes a different approach to job analysis, the basic steps associated with analyzing a job are illustrated in Fig. 1. From the perspective of job design, the writing of task statements and identifying functions of a job is one of the most important parts of a job analysis. Job functions are groups of activities that relate to a particular purpose or outcome. Figure 2 lists some typical job functions. Each job function consists of several tasks necessary to complete the job function. Tasks generally refer to an action that changes an input to an output.

From the identification of functions and tasks, the job analyst should be able to draw a fairly accurate picture of what a job entails. Nonetheless, a thorough job analysis also considers other important information that might affect performance, such as nonroutine tasks (tasks performed only occasionally), interactions with other individuals both inside and outside the organization, tasks performed as part of a team, and the position of the job within the overall organizational context.

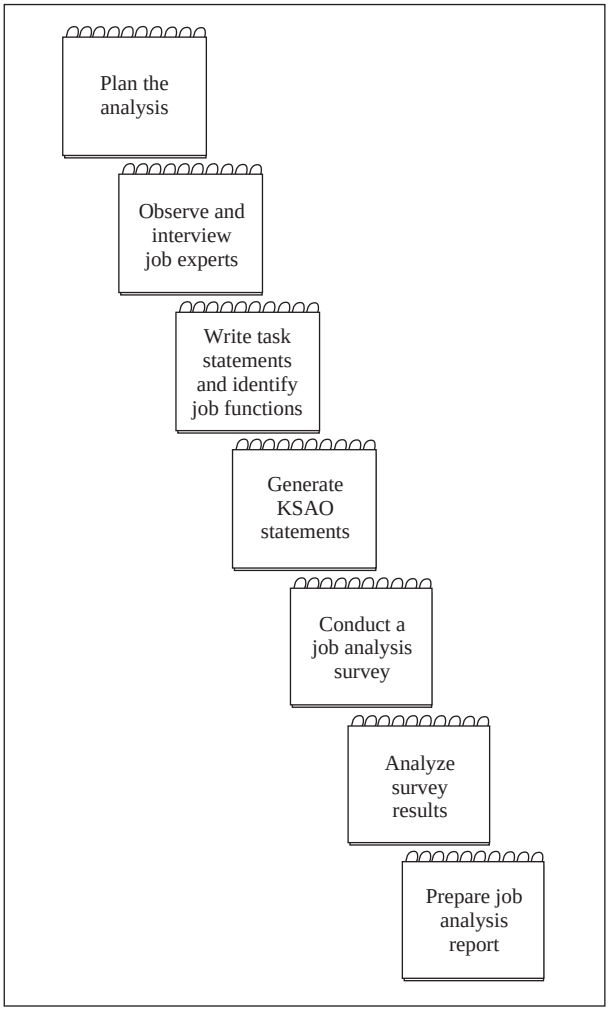

Plan the analysis
Observe and interview job experts
Write task statements and identify job functions
Generate KSAO statements
Conduct a job analysis survey
Analyze survey results
Prepare job analysis report

FIGURE 1 Steps in a job analysis. Reproduced by special permission of the publisher, Psychological Assessment Resources, Inc., 16204 North Florida Avenue, Lutz, Florida 33549, from the *Job Analysis Kit Manual*, by Sandra McIntire, Ph.D., Mary Ann Bucklan, and Deonda R. Scott, Copyright 1995 by PAR, Inc. Further reproduction is prohibited without permission of PAR, Inc.

2. JOB DESIGN

Using the information gathered from a job analysis, jobs can be designed, redesigned, created, or eliminated. Most job design starts by determining where a job fits in the overall organizational process. One way to do this is to identify the precise steps a worker is expected to perform to change an input into an output—for example, to change a consumer's application for a mortgage loan into either a commitment or denial from a mortgage lender. From this information, the job designer can identify aspects of a job aside from the specific tasks and functions identified in the job analysis. These aspects may affect performance of a job's tasks and functions. In addition, knowing each step in the process of changing an input to an output allows the job designer to evaluate the expectations associated with the job.

However, as suggested previously, focusing solely on the major functions and tasks associated with a job may not provide a complete picture of the requirements for successful performance. Other factors that may affect job performance include (i) the technology the worker would be expected to use in performing tasks and functions (e.g., machinery vs computers), (ii) the environment in which the work is performed (e.g., indoors vs outdoors or in front of customers vs producing goods for inventory), and (iii) the structure of the work organization (e.g., assembly line vs team environment).

The final part of the job design considers matching the job functions and tasks to worker abilities. That is, the job design may be affected by the abilities of workers expected to perform the job. For example, an ideal job design may call for workers to handle both production and customer service duties, but it may be unrealistic to expect workers to have high levels of ability in both these areas. In this case, the job designer would need to

Administering operations
Analyzing competition
Analyzing financial data
Buying, ordering, and stocking
Checking and testing
Cleaning and maintaining
Communicating orally
Composing or editing
Computing
Controlling expenses
Coordinating interdepartmental communications
Coping with difficulties and emergencies
Developing and implementing technology
Developing and improving onself
Developing employee potential
Developing group cooperation and teamwork
Developing sales
Filing, sorting, and routing
Financial planning and review
Gathering information
Handling outside contacts
Improving work procedures and practices
Manitaining and operating Machines/computers
Managing the floor
Merchandising and displaying products
Operating the plant
Operating the store
Processing records
Promoting community relations
Promoting safe attitudes and practices
Relating to customers
Scheduling
Setting organizational objectives
Supervising employees

FIGURE 2 Some job functions. Reproduced by special permission of the publisher, Psychological Assessment Resources, Inc., 16204 North Florida Avenue, Lutz, Florida 33549, from the *Job Analysis Kit Manual*, by Sandra McIntire, Ph.D., Mary Ann Bucklan, and Deonda R. Scott, Copyright 1995 by PAR, Inc. Further reproduction is prohibited without permission of PAR, Inc.

modify the design to reflect the practical reality of the abilities of the workers expected to perform the job.

Another consideration in job design mentioned earlier is the issue of fairness. Although it may be expedient to design jobs best suited to the mental or physical characteristics of a particular group, this kind of job design may not meet either the legal requirements or goals of the organization with regard to workforce composition. The job designer may need to modify the ideal design to reflect considerations other than performing the tasks associated with the job functions. Another important consideration in job design is worker motivation, which is discussed next.

2.1. Job Design and Worker Motivation

Three influential models of job design that consider the issue of worker motivation are the two-factor model, sociotechnical systems, and the job characteristics model.

2.1.1. The Two-Factor Model

The job simplification model developed by Taylor was later challenged by Frederick Herzberg and associates, who argued that motivational factors are a critical aspect of job design. In Herzberg's formulation, known as two-factor theory and illustrated in Fig. 3, different aspects of jobs could be categorized into what he called hygiene factors and motivators. Hygiene factors refer to the conditions of work and include salary, management, work environment, and other factors.

Motivators, on the other hand, include the tasks associated with the job as well as responsibility, autonomy, and opportunities for advancement and professional growth. According to two-factor theory, motivators have a more important impact on job performance than hygiene factors. Whereas employees may complain if they are dissatisfied with hygiene factors, their performance is more influenced by motivators. For this reason, job designers need to pay particular attention to motivators.

Herzberg's theory contradicted Taylor's belief that simplification should be the basis for designing jobs. By recognizing that performance is affected by factors other than tasks, Herzberg advocated an approach known as job enrichment. Job enrichment refers to designing jobs so that workers have more responsibility for their accomplishment. In this way, workers are more motivated to accomplish their jobs at a higher level of performance. Job enlargement is a kind of enrichment that attempts to make jobs more interesting by adding duties rather than adding responsibility.

Although studies of the two-factor model indicate that motivation and performance are more complex



FIGURE 3 Motivators and hygiene factors. *Because of its ubiquitous nature, salary commonly shows up as a motivator as well as a hygiene factor. Although primarily a hygiene factor, it also often takes on some of the properties of a motivator, with dynamics similar to those of recognition for achievement. From Herzberg, F. (1982). *The managerial choice: To be efficient and to be human* (2nd rev. ed.). Salt Lake City, UT: Olympus.

than the model suggests, Herzberg's theory had significant intuitive appeal when applied to understanding worker motivation. From the perspective of job design, a shortcoming of two-factor theory is its lack of specificity in identifying how job tasks become motivators that affect performance.

2.1.2. Sociotechnical Systems

A second important influence on job design was developed at the Tavistock Institute of Human Relations in England. This model, developed during the period after World War II as Europeans introduced technology into their factories, argued that organizations consist of two systems—the social and the technical—that affect worker performance. According to sociotechnical systems theory, jobs should be designed with both their social and their technical aspects in mind.

Sociotechnical theorists believe that introducing technological changes into a workplace almost always disrupts the social system. These disruptions, in turn, are likely to lower worker satisfaction and productivity. Therefore, a job design that does not consider the social aspects of a job is unlikely to result in optimal performance.

An important aspect of the sociotechnical approach to job design—and an important difference from job simplification—is its emphasis on work groups rather than individual jobs. Sociotechnical theorists believe that designing jobs in terms of work groups that have autonomy to plan how to accomplish their tasks results in both greater productivity and greater worker satisfaction. They also argue that increases in productivity

occur more often if workers develop multiple skills and if jobs have a variety of duties, which is a kind of job enrichment.

2.1.3. The Job Characteristics Model

Probably the most influential approach used by modern job designers is the job characteristics model developed by J. R. Hackman and Greg Oldham. Although Hackman and Oldham's research focused on worker motivation, their identification of five domains associated with any job has been useful in the study of job design. Figure 4 illustrates the job characteristics model.

Skill variety refers to the abilities a worker uses to accomplish a job. According to the job characteristics model, jobs that require a wider range of skills are more motivating. Task identity refers to the degree to which the worker sees the outcome of his or her efforts, and task significance relates to the impact of the job on others, both inside and outside the organization. According to the theorists, the three aspects of skill variety, task identity, and task significance affect the worker's assessment of the importance of his or her job.

The other two factors in the model address the responsibility associated with the job and the worker's perception of outcomes. Autonomy refers to the level of independence the worker has in deciding how to accomplish the tasks associated with the job, and feedback describes the degree to which the worker receives information about his or her performance.

A final aspect of the job characteristics model addresses an important individual difference variable that affects job performance. Growth need strength refers to the worker's need for personal growth through job performance. Growth need strength can affect job design in that workers with high growth need strength seek jobs that offer challenge and a variety of duties; workers with low growth need strength prefer more routine types of jobs. If a job designer is dealing with jobs requiring a highly educated workforce, for example, growth need strength may require that jobs be made more complex and that levels of autonomy be higher.

3. JOB EVALUATION

As suggested previously, job evaluation is the procedure used to determine the relative worth of jobs within an organization. Most jobs are evaluated in terms of education required, skills involved in job performance, level of responsibility, working conditions, and factors in the local economy. Information from the evaluation is used to determine salary levels for jobs.

There are several methods for job evaluation. The point method is the most widely used. Under this procedure, subject matter experts identify job dimensions

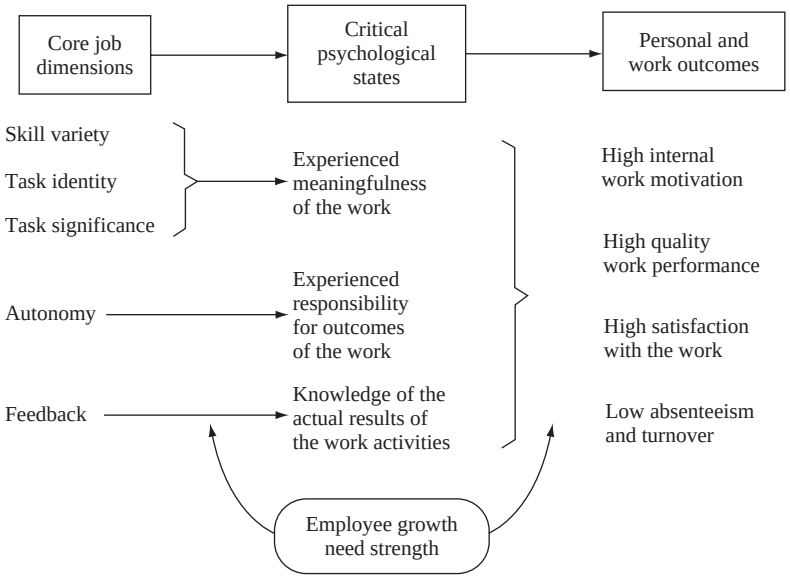


FIGURE 4 Job characteristics model. Reprinted from *Organizational Behavior and Human Performance, 16*, Hackman, J. R., & Oldham, G. R., Motivation through the design of work: Test of a theory, 250–279, Copyright 1976, with permission from Elsevier.

that will be the basis for assigning salary level. These dimensions—known as commensurable factors—are then described in terms of behaviors typical of different levels of skill and assigned a number that reflects differences in skill levels. Jobs can then be scored in terms of the number of points and compared to each other to determine salary level.

Job ranking is the most straightforward method of job evaluation and is most often used in small organizations. In this procedure, a SME ranks all jobs in an organization or a department in terms of the importance of their contribution to the overall organizational goals. In contrast to the point method, job ranking is a subjective procedure that does not consider the magnitude of the difference between jobs.

Finally, the job grading method places jobs into levels or grades. Jobs requiring less skill, for example, would be placed together—irrespective of the work performed—in a lower grade, whereas jobs requiring more skill would be grouped in a higher grade. The U.S. government's GS system is an example of job grading.

Regardless of the procedure used, job evaluation results in a hierarchical ranking of the worth of each job within the organization. The final step in determining salaries is external validation of pay scales. This is usually done through surveys of salaries paid for similar jobs in other organizations.

4. OTHER ISSUES IN JOB ANALYSIS, DESIGN, AND EVALUATION

Because of the complexities of modern organizations, considerations other than functions and tasks affect the way jobs are structured and salary levels assigned. Some issues that may influence the way an organization analyzes, designs, and evaluates jobs include individual differences, organization development, and teams.

4.1. Individual Differences

When Taylor first starting designing jobs at Bethlehem Steel, his goal was to standardize procedures for job performance so that any person could do the work. In general, standardization remains the basic principle of job design, which is that anyone with the requisite KSAs should be able to do the work.

However, as mentioned previously, job performance is greatly influenced by level of motivation, which is an individual difference variable. When motivation to perform is high, jobs can be designed that encompass a variety of tasks and challenge the worker. When motivation is lower, jobs may need to be simpler, encompassing fewer tasks. Similarly, ability may also impact a worker's motivation and performance.

Another relevant individual difference variable is the tendency of individual workers to redesign their jobs so that they spend more time performing duties they enjoy and less time performing duties they find unappealing. In these cases, the job design provides a general outline for performance rather than a strict guideline. An employee may redesign a job so that he or she spends more or less time on certain tasks and still maintain a high level of performance overall.

4.2. Organization Development

During the 1990s, organizational theorists increasingly argued that the organization of the future would emphasize flexibility over structure. This argument was based on a variety of factors, but among the most important were changes in workforce demographics, technological innovations, and globalization.

As the U.S. workforce widened to include individuals from different races and cultural backgrounds, some theorists argued that traditional models of workplace organization were not appropriate for everyone. In particular, U.S. managerial practices were criticized for not taking differences such as race, culture, and gender into consideration. These criticisms, in turn, gave rise to the diversity movement. Diversity considerations may not affect the design of every job, but, at a minimum, they may be a factor in determining the physical requirements for performing the tasks associated with a job. For example, women, as a group, are not as physically strong as men, and Asians, as a group, are not usually as large as Westerners. Changes in the composition of the workforce may require job designers to consider the type of person most likely to do the job being designed as well as the tasks and functions associated with the job.

Technology is also likely to change job design. Although technology minimizes certain differences between workers (e.g., physical strength), technology often requires specific skills, such as facility with computers and computer programs, vigilance, and an ability to grasp the underlying model under which technology operates. A related challenge to traditional job design and evaluation is the changing nature of technology. Because technology changes rapidly in

some occupations, constant revision of the design and requirements of specific jobs may be necessary.

Finally, globalization affects job design in a variety of ways. Aside from confronting cultural differences, workers with foreign colleagues or who produce products for foreign markets may encounter factors such as language differences, time differences, incompatibilities in technology, or even political considerations that affect the way jobs are designed and evaluated.

4.3. Teams

During the 1980s, organizations began to pay more attention to the role of teams in the work environment. Although sociotechnical theorists had advocated teamwork as a means of both making work more meaningful to the worker and enhancing productivity since the 1950s, only a few U.S. organizations had adopted a teamwork approach. Building on an interest in Japanese quality circles that emerged in the 1970s, U.S. managers began to adapt team concepts to their workforces. In the modern team approach, teams often have responsibility for determining work schedules, methods, and hours, as well as team leadership and communication procedures, all of which can affect job design.

Teamwork calls for a radically different approach to job design. Rather than focus on requirements associated with a specific job, job design for teams requires the analysis of several jobs concurrently as well as the establishment of requirements based on the relationship of each job to other jobs represented by team members. This can be a complicated procedure, and most approaches to job design have not addressed the question of interrelationships among jobs held by members of the same team. Similarly, organizing workers into teams responsible for accomplishing a particular task may complicate the job evaluation process.

See Also the Following Articles

Motives and Goals ■ Organization Development ■ Work Motivation ■ Work Teams

Further Reading

Davis, L. E., & Wacker, G. J. (1988). Job design. In S. Gael (Ed.), *The job analysis handbook for business, industry, and government* (Vol. 1). New York: Wiley.

Hackman, J. R., & Oldham, G. R. (1976). Motivation through the design of work: Test of a theory. *Organizational Behavior and Human Performance*, *16*, 250–279.

Herzberg, F., Mausner, B., & Snyderman, B. (1959). *The motivation to work*. New York: Wiley.

McIntire, S. A., Bucklan, M. A., & Scott, D. R. (1995). *Job analysis kit*. Odessa, FL: Psychological Assessment Resources.

Parker, S. K., & Wall, T. D. (2001). In N. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswesvaran (Eds.), *Handbook of industrial, work, and organizational psychology* (Vol. 1). London: Sage.

Smither, R. D. (1988). *The psychology of work and human performance*. New York: Harper & Row.

Trist, E. L. (1981). The evolution of sociotechnical systems as a conceptual framework and as an action research program. In A. H. Van de Ven, & W. F. Joyce (Eds.), *Perspectives on organization and behavior*. New York: Wiley.

Job Search

Alan M. Saks
University of Toronto, Toronto, Ontario, Canada

1. The Job Search Process
2. Job Search Research
3. Job Search Behaviors
4. Antecedents of Job Search Behaviors
5. Consequences of Job Search Behaviors
6. Conclusion
Further Reading

GLOSSARY

active job search behavior Applying for positions by sending out resumes and interviewing with prospective employers.
assertive job seeking The ability to identify one's rights and choices during job search and to act on them.
coping resources Individual and environmental factors and conditions that an individual can draw on in coping with involuntary job loss and that have direct and indirect effects on coping strategies.
emotion-focused strategies Attempts to change the way that the stressor is construed through techniques such as distancing and avoidance.
emotional response model Job seekers experience high levels of stress and frustration that can lead to avoidance, helplessness, and withdrawal.
employment commitment The importance and centrality of employment to the job seeker.
employment outcome Whether or not a job seeker obtains employment and the nature of that employment.
employment quality Job search outcomes that occur once the job seeker assumes a position and begins working, such as job satisfaction and intentions to quit.
employment status Whether or not a job seeker has obtained employment.
formal sources Public intermediaries such as advertisements, employment agencies, and campus placement offices.
informal sources Private intermediaries such as friends, relatives, or persons who are already employed in an organization.
job-finding club An intensive, highly structured, group-based behavioral counseling program that emphasizes motivation, maintenance of behavior, feedback and reinforcement, imitation, role playing, and practice activities.
job information sources The means by which job seekers learn about job opportunities.
job search A dynamic process through which individuals pursue employment goals.
job search effort The amount of energy, time, and persistence that a job seeker devotes to his or her job search.
job search intensity The frequency with which job seekers, during a set period of time, engage in specific job search behaviors or activities, such as preparing a resume or contacting an employment agency.
job search outcomes The immediate outcomes that occur during one's job search and lead to subsequent employment outcomes.
job search self-efficacy Confidence in one's ability to perform a variety of specific job search tasks and activities.
job-seeking social support The network of friends and family who provide counseling, assistance, and encouragement to job seekers.
JOBS intervention A job search intervention that includes problem-solving and decision-making processes, inoculation against setbacks, social support, and job-seeking skills.

©2004 Elsevier Inc.
All rights reserved.

learning model Job seekers learn more efficient and effective search techniques during their job search.

networking Contacting friends, acquaintances, and referrals to obtain information and leads about job opportunities.

preparatory job search behavior Gathering job search information and identifying potential leads during the planning phase of job search.

problem-focused coping strategies Activities directly aimed at eliminating the source of stress (i.e., job loss) such as proactive job search.

sequential model Job search follows a logical and systematic progression of stages.

symptom-focused coping attempts Attempts to eliminate the symptoms of stress through involvement in nonwork activities or seeking social support from friends and family.

Job search is a dynamic process through which individuals pursue employment goals. Whether one is entering the workforce for the first time, seeking reemployment following job loss, or changing jobs to pursue new career opportunities, job search is a necessary and inevitable part of the process. Numerous individual difference variables and situational factors combine to shape a person's job search. In turn, the nature and scope of an individual's job search can influence the probability and speed that one will obtain employment. For unemployed job seekers, job search is a form of problem-focused coping in response to job loss. An effective job search can increase an individual's chances of reemployment following job loss and improve long-term psychological well-being. Among employed job seekers, job search is a major predictor of voluntary turnover. The extent to which one engages in an effective job search and finds satisfactory employment can be enhanced through job search training interventions.

1. THE JOB SEARCH PROCESS

Job search is a process that consists of gathering information about potential job opportunities, generating and evaluating job alternatives, and choosing a job from the alternatives. These activities determine the type and amount of information that job seekers obtain about job openings as well as the number of job opportunities from which a job seeker has to choose.

The job search process has often been conceived of as a logical sequence of activities. Soelberg's model of job search consists of two phases: planning job search and then job search and choice. Thus, job search begins with an extensive search to gather information and identify job opportunities followed by a more intensive search that involves the acquisition of specific information about jobs and organizations. Job search has also been described as consisting of a preparatory and active phase. Preparatory job search behavior involves information gathering about job opportunities, and active job search behavior involves applying for positions.

Kanfer *et al*. conceptualized job search as a self-regulatory process that involves "a purposive, volitional pattern of action that begins with the identification and commitment to pursuing an employment goal." Employment goals activate job search behaviors that are intended to obtain employment. Thus, "individuals identify, initiate, and pursue actions for the purpose of obtaining new employment or reemployment." Job search and the self-regulatory process ends when the employment goal has been achieved or is abandoned.

Finally, it is worth noting that during the search process job seekers change their behaviors. Three models describe how the behavior of job seekers can change during their job search. As indicated previously, the sequential model proposes that job search follows a logical and systematic progression of stages. According to the learning model, job seekers learn more efficient and effective search techniques during their job search. The emotional response model asserts that job seekers experience high levels of stress and frustration that can lead to avoidance, helplessness, and withdrawal.

Two studies have examined changes in job seeker's search behavior. Barber and colleagues measured the job search activities of college and vocational–technical school graduates early in the search process, at graduation, and 3 months following graduation for those who remained unemployed. They found a significant decrease from initial search to late search in the use of formal sources, job search intensity, and in information related to obtaining a job. They also reported a significant increase in the use of formal sources and intensity following late search for unemployed job seekers.

Saks and Ashforth studied recent university graduates who had not found a job just prior to graduation. Follow-up 4 months later indicated that job seekers increased their active job search behavior, formal job source usage, and search effort and decreased their job search anxiety.

2. JOB SEARCH RESEARCH

Research on job search can be distinguished in terms of the major area or domain of interest as well as the sample. Job search research has tended to fall into one of the following areas: (i) coping with job loss and reemployment, (ii) job search interventions, (iii) job search and voluntary turnover, and (iv) antecedents and consequences of job search.

The samples used in job search research have varied across and within the main research areas. For example, research on coping with job loss and reemployment necessarily includes unemployed and recently laid-off workers. Research on job search and turnover includes samples of employed job seekers. Research on the antecedents and consequences of job search has included new entrants to the workforce (e.g., college graduates), employed job seekers, as well as unemployed job seekers. However, it is worth noting that college students have been used most extensively in job search research, followed by unemployed job seekers.

2.1. Coping with Job Loss and Reemployment

Studies on job search and involuntary job loss focus on the coping strategies of recently unemployed individuals. Job loss coping strategies can be problem-focused, emotion-focused, and symptom-focused. Problem-focused coping strategies involve activities directly aimed at eliminating the source of stress (i.e., job loss) such as proactive job search. Emotion-focused strategies involve attempts to change the way that the stressor is construed through techniques such as distancing and avoidance. Symptom-focused coping attempts to eliminate the symptoms of stress through involvement in nonwork activities or seeking social support from friends and family.

Research on coping with job loss measures the predictors and consequences of coping strategies. Predictors of coping strategies consist of a number of coping resources, including individual and environmental factors and conditions that an individual can draw on in coping with involuntary job loss and that have direct and indirect effects on coping strategies. Coping resources such as education, financial resources, and social support predict cognitive appraisal of job loss and coping strategies, including job search activities. Reemployment coping goals and gender also predict proactive job search. Men engage in more problem-focused coping strategies such as job search, whereas women engage in more symptom-focused strategies such as social support.

The consequences of coping behaviors include mental health and reemployment. Problem-focused coping (i.e., proactive job search) relates to obtaining reemployment but can also reduce mental health in the short term.

2.2. Job Search Interventions

Job search interventions can lessen the negative effects of job loss and improve individuals' job search and probability of reemployment. Research on employment counseling and job loss has found that the job-finding club is one of the most effective methods of job search training.

The job-finding club is an intensive, highly structured, group-based behavioral counseling program that emphasizes motivation, maintenance of behavior, feedback and reinforcement, imitation, role playing, and practice activities. Participants receive assistance in all areas of job search, including coping with discouragement, preparing resumes, obtaining and pursuing job leads, learning interviewing skills, scheduling time and record keeping, dress and grooming, and using the telephone for making inquiries and contacts.

Job seekers who have attended a job-finding club have been found to be more likely to obtain employment compared to individuals in a control group or who participated in other types of programs. Some studies have also found the job-finding club to result in higher starting salaries, more work hours, lower depression, and greater advancement and job satisfaction.

A number of other job search interventions have also been shown to be effective for coping with job loss and reemployment. For example, Caplan and colleagues developed a program called the JOBS intervention to improve the motivation and skills necessary for unemployed job seekers to engage in job-seeking behavior. The intervention includes problem-solving and decision-making processes, inoculation against setbacks, social support, and job-seeking skills.

Compared to a control group, the JOBS intervention resulted in higher rates of reemployment as well as higher quality reemployment in terms of earnings and job satisfaction for its participants. For those who remained unemployed, participants in the JOBS program have higher job search motivation and self-efficacy compared to control group participants. The benefits of the intervention for reemployment and employment quality have been shown to persist. Individuals who have participated in the JOBS program have higher levels

of employment, higher paying jobs, work more hours, and had fewer changes in jobs and employers up to 2½ years following the program. The JOBS intervention also has a significant effect on the mental health of individuals at high risk for depression.

Eden and Aviram studied the effects of a behavioral modeling job search workshop designed to boost the general self-efficacy (GSE) of unemployed vocational workers. Participants who attended the workshop had higher self-efficacy at the end the workshop and 2 months later compared to those in a control group. The workshop also increased the job search activity and reemployment of individuals with low initial GSE. Furthermore, the effect of the workshop on reemployment was mediated by job search activity.

Finally, a disclosive writing intervention in which laid-off professionals wrote about their thoughts and feelings about the layoff and how it affected their personal and professional lives resulted in higher levels of full-time employment compared to those for a control group.

2.3. Job Search and Voluntary Turnover

Models of turnover and the turnover process often include job search as a key predictor of voluntary turnover. Several studies have found that job search is highly predictive of turnover and even more strongly related to turnover than more traditional predictors such as perceptual, affective, attitudinal, and intention measures.

In terms of specific job search variables, Gary Blau has shown that active job search behavior is a better predictor of voluntary turnover than preparatory job search behavior and general search effort, and that preparatory job search behavior is indirectly related to turnover through active job search behavior. In addition, active job search behavior predicts voluntary turnover beyond job attitudes and withdrawal cognitions. Job search has also been found to predict the voluntary turnover of managers and to mediate the relationship between motivation variables and turnover.

Thus, it appears that job search, rather than turnover intentions, is the best predictor of voluntary turnover. Although this stream of research does not measure the employment status of those who quit, it is safe to assume that if an employee has quit he or she has found another job. Therefore, the finding that job search is related to turnover is akin to stating that job search is related to finding a new job.

2.4. Antecedents and Consequences of Job Search

In 1987, Schwab and colleagues presented a job search model in which self-esteem and financial need predicted job search intensity, and job search intensity predicted employment. Since then, additional antecedents, job search variables, and outcomes have been investigated. Given that this research represents the most extensive area of job search and overlaps with the other areas, the following sections focus on job search behaviors, antecedents, and consequences.

As indicated previously, research on the antecedents and consequences of job search has varied as a function of sample type (i.e., new entrants, employed, and unemployed job seekers). Since the characteristics of these samples differ in terms of experience, motivation, and the use of job search strategies, the antecedents and consequences of job search might differ across studies and samples. In fact, in a meta-analysis on the antecedents and consequences of job search conducted by Kanfer and colleagues, sample type was a significant moderator variable.

3. JOB SEARCH BEHAVIORS

In their review of the job search literature, Schwab and colleagues indicated that job search success and employment outcomes are a function of the sources used to acquire information about job openings and the intensity with which one pursues such information. Although sources and intensity are among the most often studied job search dimensions, job search research has included a number of other dimensions of job search, the four most prominent being job information sources, job search intensity, job search effort, and job search information.

3.1. Job Information Sources

Job information sources is considered to be one of the most important dimensions of job search and is among the most often studied. Sources refer to the means by which job seekers learn about job opportunities. Both the recruitment and the job search literature have tended to distinguish between formal and informal job sources. Formal sources involve the use of public intermediaries, such as advertisements, employment agencies, and campus placement offices. Informal sources are private intermediaries, such as friends,

relatives, or persons who are already employed in an organization.

3.2. Job Search Intensity

Job search intensity refers to the frequency with which job seekers, during a set period of time, engage in specific job search behaviors or activities, such as preparing a resume or contacting an employment agency. Various indicators have been used to measure search intensity, such as the number of employers contacted, the number of hours per week spent searching, and the number of sources used.

In response to the failure to distinguish between job search activities at different stages of the job search process as well as the frequent use of single-item measures and dichotomous response scales, Gary Blau developed two distinct measures of job search intensity that he called preparatory job search behavior and active job search behavior. Preparatory job search behavior involves gathering job search information and identifying potential leads during the planning phase of job search. Active job search behavior involves the actual job search and choice process, such as sending out resumes and interviewing with prospective employers. It reflects an individual's behavioral commitment to job search.

Several studies have measured specific job search behaviors that have been recommended in the popular job search literature, particularly assertiveness and networking. Assertive job-seeking applies the concept of assertiveness to job search and refers to the ability to identify one's rights and choices during job search and to act on them. Specific behaviors include making follow-up calls regarding the status of a job application and making calls to arrange meetings with organizational representatives to discuss employment opportunities. Networking is another specific job search behavior and involves contacting friends, acquaintances, and referrals to obtain information and leads about job opportunities. Networking involves the use of informal job information sources.

3.3. Job Search Effort

Job search effort refers to the amount of energy, time, and persistence that a job seeker devotes to his or her job search. Unlike measures of job search intensity, they do not focus on particular job search behaviors. Blau developed a scale to measure general effort in job search that asks respondents about the time and effort they have devoted to their search. Based on confirmatory factor analyses across there diverse samples, he found empirical support that general effort is related to, but distinguishable from, preparatory and active job search behavior.

3.4. Job Search Information

Barber and colleagues argued that the type of information sought and acquired is also an important dimension of job search because individuals use this information to make job choice decisions. In addition to information about employment opportunities, job seekers also require accurate, detailed, and specific information about job and organizational attributes. Few studies, however, have actually examined this dimension of job search.

4. ANTECEDENTS OF JOB SEARCH BEHAVIORS

A number of variables have been found to influence job search and employment outcomes. In general, the antecedents of job search can be categorized into three groups: biographical/demographic variables, individual difference variables, and situational factors.

However, the size of the relationships between the antecedents and job search behavior has been found to differ as a function of the job search measure and the sample type. In a meta-analysis, Kanfer and colleagues concluded that a number of antecedents–job search relationships were stronger for measures of job search intensity compared to measures of job search effort. Also, for some antecedents (i.e., extroversion, neuroticism, conscientiousness, and education) the relationships were stronger for new entrants, whereas for other antecedents (i.e., self-esteem, employment commitment, and social support) the relationships were stronger for unemployed job seekers. In addition, tenure related positively to search behavior for new entrants but related negatively to search behavior for employed and unemployed job seekers. It is also worth noting that the relationships between the antecedents and search behavior tend to be stronger than the relationships between the antecedents and employment outcomes.

4.1. Demographic/Biographical Variables

Demographic/biographical variables in job search research include age, gender, education, race, and tenure.

In general, biographical variables tend to be only weakly related to job search behavior and to a lesser extent than the other antecedents. In their meta-analysis, Kanfer and colleagues concluded that age, race, gender, education, and tenure were related to job search behavior; that is, older, nonwhite, female, less educated, and longer tenured individuals engaged in less job search behavior than younger, white, male, more educated, and less tenured individuals. Furthermore, men have been found to be more likely to use informal sources than women.

Demographic variables are also related to job search outcomes. Younger job seekers as well as more educated ones and males (new entrants only) are more likely to find employment, and more educated and white job seekers experience a shorter period of unemployment.

Finally, there is evidence that cognitive ability may also play a role in job search. For example, the cognitive ability of employed managers has been shown to relate positively to job search. Also, grade average, a surrogate measure of cognitive ability, has been found to correlate positively to active and preparatory job search behavior as well as employment status among a sample of recent graduates.

4.2. Individual Difference Variables

Individual difference variables tend to be personality variables (e.g., self-esteem), motivational factors (e.g., self-efficacy), and attitudes toward employment and work (e.g., employment commitment). Among the personality variables, self-esteem has been found to relate negatively to the use of formal job sources and relate positively to job search intensity and assertive job-seeking behavior, as well as a shorter period of unemployment, the number of job offers received, and the likelihood of obtaining employment.

There is evidence that the dimensions of the five-factor model of personality are related to job search behaviors. In their meta-analysis, Kanfer and colleagues organized personality constructs used in job search research on the basis of the five-factor model of personality. They found that extroversion and conscientiousness were the strongest positive predictors of job search behavior, followed by openness to experience and agreeableness. Neuroticism related negatively to job search behavior. In addition, higher levels of extroversion, conscientiousness, openness to experience, and agreeableness related to a shorter period of unemployment, and conscientiousness was also related to employment status.

A motivational variable that has received a considerable amount of attention is job search self-efficacy—that is, confidence in one's ability to perform a variety of specific job search tasks and activities. The results of several studies indicate that job search self-efficacy predicts preparatory and active job search behaviors, assertive job search behavior, job source usage, and search intensity, as well as search duration, number of job offers, and employment status.

Another motivational variable that has been shown to be important is job seekers' perceived control over the outcomes of their job search. Although locus of control has been found to relate only weakly to job search behavior, perceived situational control over the outcomes of job search has been found to relate positively to job search behaviors and employment status.

In terms of employment attitudes, employment commitment or the importance and centrality of employment to the job seeker has been found to relate positively to job search intensity. Among employed job seekers, job attitudes such as job satisfaction and organizational commitment have been found to relate negatively to job search behavior.

4.3. Situational Variables

Among the situational variables, financial need and social support have received a considerable amount of attention. Job seekers who have a greater financial need tend to be more intense in their search for employment. Financial need has been shown to relate positively to both preparatory and active job search behaviors. Job seekers with higher benefit levels or a longer duration of benefits tend to remain unemployed for longer periods. Thus, financial hardship tends to result in a shorter period of unemployment.

One of the most important situational variables in job search and coping with job loss is social or job-seeking support. Social support refers to the network of friends and family who provide counseling, assistance, and encouragement to job seekers. Social support has been shown to be particularly important in assisting individuals following job loss and to predict job search behavior and employment status.

Finally, several other situational factors have been found to predict the job search behavior of employed individuals. In particular, compensation, perceived organizational success, and organizational work–family balance policies have been found to relate negatively to job search. Job insecurity and desire for more work–family balance have been found to relate positively to

job search. These results are particularly noteworthy because they suggest that organizations have some influence over a number of job and organizational variables that affect employees' job search behavior.

5. CONSEQUENCES OF JOB SEARCH BEHAVIORS

In their model of job search, Schwab and colleagues proposed two outcomes of job search: employment and the quality of employment (e.g., job satisfaction and tenure). Job search research, however, has focused primarily on preentry outcomes, especially employment status (i.e., whether or not the job seeker has obtained employment). In general, the consequences of job search can be categorized in terms of (i) job search outcomes (outcomes that occur during the search process), (ii) employment outcomes (outcomes that are a result of one's job search), (iii) psychological well-being (the job seeker's mental health), and (iv) employment quality (outcomes that occur on-the-job).

5.1. Job Search Outcomes

Job search outcomes refer to the immediate outcomes that occur during one's job search and lead to subsequent employment outcomes. This includes the number of job interviews and job offers that a job seeker receives as well as the speed with which one obtains employment. Research supports a positive relationship between job search intensity and the number of job interviews and offers received and a negative relationship with search duration.

5.2. Employment Outcomes

Employment outcome measures assess the result of one's job search and refer to whether or not a job seeker obtains employment and the nature of that employment. The most common outcome measure of job search is employment status. There is strong support for a positive relationship between job search behavior and employment status. Furthermore, as indicated previously, research showing a relationship between job search intensity and turnover provides additional support that job search results in subsequent employment. Job search behavior has also been found to relate positively to preentry person–job (P–J) fit and person–organization (P–O) fit and the initial compensation of recent graduates.

Interestingly, Kanfer and colleagues found that measures of job search effort related more strongly to employment duration and employment status compared to measures of job search intensity. In addition, the relationship between job search behavior and employment status was strongest for employed job seekers, followed by new entrants and unemployed job seekers.

Research on job sources has consistently found that job seekers are more likely to find employment through the use of informal sources of information. In particular, friends and personal acquaintances tend to be the main source through which job seekers obtain employment.

However, although informal sources may be more likely to result in employment, there is evidence that formal sources are more likely to result in a better job. For example, graduates who used formal sources have been found to be more likely to find work in higher level occupations and jobs that are more closely related to their college training than those who used informal sources. Also, the number of formal sources used by recent graduates has been found to be positively related to preentry perceptions of P–J and P–O fit. Formal sources have also been found to result in higher paying jobs.

5.3. Psychological Well-Being

A major concern in the job search literature is the psychological well-being of job seekers. This is especially the case in research on job loss, unemployment, and job search interventions. It is well known that the experience of job loss and unemployment is associated with negative physical and psychological effects, including symptoms of psychiatric disorder, distress, and depression.

Unlike the other outcomes discussed in this section, indicators of psychological well-being tend not to be directly associated with job search behavior but, rather, the experience of (un)employment. In this regard, job seekers who obtain satisfactory employment tend to show an improvement in their mental health and psychological well-being. However, given that job search is related to (re)employment, it is fair to say that job search behavior is indirectly related to psychological well-being through its influence on employment status.

Finally, job search interventions have also been shown to influence job seekers' psychological well-being. A number of such interventions have improved unemployed job seekers' self-esteem, self-efficacy, depression, and mental health.

5.4. Employment Quality

Employment quality refers to job search outcomes that occur once the job seeker assumes a position and begins working. With the exception of research on recruitment and job sources, job search research has seldom measured the quality of employment. Research on unemployment has also tended to neglect the quality of reemployment.

The relationship between recruitment sources and postentry outcomes has been extensively studied in the recruitment literature. Accordingly, job seekers who find employment through informal sources tend to have more positive job attitudes and lower turnover.

Job search research, however, has only recently begun to examine the links between job search behavior and employment quality, and the results have not been very supportive. For example, Wanberg and colleagues did not find a significant relationship between job search intensity and job satisfaction, job improvement, and intentions to turnover. In another study, they found that neither job search nor networking intensity related to job satisfaction or intention to turnover. Werbel also failed to detect a significant relationship between job search intensity and job satisfaction. Furthermore, research on involuntary job loss has also failed to detect a significant relationship between problem-focused coping and the quality of reemployment.

However, there is evidence that job search behavior is indirectly related to employment quality. Saks and Ashforth found that job search behavior was indirectly related to job and organizational attitudes through a direct relationship with preentry perceptions of P–J and P–O fit. In a related study, they found that preentry fit perceptions mediated the relationship between job information sources and job satisfaction, intentions to quit, and turnover.

6. CONCLUSION

Job search is a major factor in the experience of job loss, turnover, psychological well-being, and (re)employment. Both individual differences and situational factors influence job search behavior, and the nature and scope of one's job search are related to the probability, speed, and quality of employment that one obtains. Given that workers today are expected to experience more than a dozen job transitions during the course of their careers, the need for individuals to have job search skills is critical for job seekers' continued employment and psychological well-being. Fortunately, the knowledge exists for career counselors to assist job seekers in coping with job loss and finding satisfactory (re)employment.

See Also the Following Articles

Coping ■ Employment Interviewing ■ Job Analysis, Design, and Evaluation ■ Occupational Choice ■ Occupational Psychology, Overview ■ Social Networks

Further Reading

Barber, A. E., Daly, C. L., Giannantonio, C. M., & Philips, J. M. (1994). Job search activities: An examination of changes over time. *Personnel Psychology*, *47*, 739–765.

Blau, G. (1993). Further exploring the relationship between job search and voluntary individual turnover. *Personnel Psychology*, *46*, 213–330.

Blau, G. (1994). Testing a two-dimensional measure of job search behavior. *Organizational Behavior and Human Decision Processes*, *59*, 288–312.

Caplan, R. D., Vinokur, A. D., Price, R. H., & van Ryn, M. (1989). Job seeking, reemployment, and mental health: A randomized field experiment in coping with job loss. *Journal of Applied Psychology*, *74*, 759–769.

Eden, D., & Aviram, A. (1993). Self-efficacy training to speed reemployment: Helping people to help themselves. *Journal of Applied Psychology*, *78*, 352–360.

Granovetter, M. S. (1995). *Getting a job* (2nd ed.). Chicago: University of Chicago Press.

Hanisch, K. A. (1999). Job loss and unemployment research from 1994 to 1998: A review and recommendations for research and intervention. *Journal of Vocational Behavior*, *55*, 188–220.

Kanfer, R., Wanberg, C. R., & Kantrowitz, T. M. (2001). Job search and employment: A personality-motivational analysis and meta-analytic review. *Journal of Applied Psychology*, *86*, 837–855.

Latack, J. C., Kinicki, A. J., & Prussia, G. E. (1995). An integrative process model of coping with job loss. *Academy of Management Review*, *20*, 311–342.

Leana, C. R., & Feldman, D. D. (1992). *Coping with job loss*. New York: Lexington.

Saks, A. M., & Ashforth, B. E. (1997). A longitudinal investigation of the relationships between job information sources, applicant perceptions of fit, and work outcomes. *Personnel Psychology*, *50*, 395–426.

Saks, A. M., & Ashforth, B. E. (2000). Change in job search behaviors and employment outcomes. *Journal of Vocational Behavior*, *56*, 277–287.

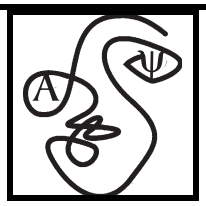
Saks, A. M., & Ashforth, B. E. (2002). Is job search related to employment quality? It all depends on the fit. *Journal of Applied Psychology*, *87*, 646–654.

Schwab, D. P., Rynes, S. L., & Aldag, R. J. (1987). Theories and research on job search and choice. In K. M. Rowland, & G. R. Ferris (Eds.), *Research in personnel and human resources management* (Vol. 5, pp. 129–166). Greenwich, CT: JAI Press.

Soelberg, P. O. (1967). Unprogrammed decision making. *Industrial Management Review*, *8*, 19–29.

Wanberg, C. R. (1997). Antecedents and outcomes of coping behaviors among unemployed and reemployed individuals. *Journal of Applied Psychology*, *82*, 731–744.

Wanberg, C. R., Kanfer, R., & Rotundo, M. (1999). Unemployed individuals: Motives, job search competencies, and job search constraints as predictors of job seeking and reemployment. *Journal of Applied Psychology*, *84*, 897–910.



Job Stress

James Campbell Quick and Marilyn Macik-Frey
University of Texas, Arlington, Texas, USA

Debra L. Nelson
Oklahoma State University, Stillwater, Oklahoma, USA

1. Introduction
2. Causes of Job Stress
3. Consequences of Job Stress
4. Modifiers of the Stress Response
5. Measures of Job Stress
6. Early Warning Signs of Strain
7. Preventive Stress Management
8. Implications for Executives and Leaders
Further Reading

GLOSSARY

eustress The healthy, positive, constructive outcome of stressful events and the stress response.
job stress The mind–body arousal resulting from physical and/or psychological stressors associated with a job and work.
preventive stress management An approach and set of principles for promoting eustress and health while preventing strain and dysfunction.
primary prevention The modification, reduction, or management of sources of stress.
secondary prevention The modification, reduction, or management of the individual's stress response.
strain The physiological, psychological, and/or behavioral deviation from an individual's healthy functioning.
stress response The generalized, patterned, unconscious mobilization of the body's natural energy and resources.
stressor A physical or psychological stimulus that causes the individual stress response.
tertiary prevention The treatment or caregiving intervention to heal a person's strain.
trauma A highly stressful, deeply disturbing, or dramatically unexpected stressful event experienced as traumatic by the individual.

Job stress is the psychophysiological arousal resulting from workplace demands. Optimally, it results in enhanced job performance. When mismanaged, job stress leads to job strain, which includes psychological, medical, and behavioral costs.

1. INTRODUCTION

Job stress is important for two reasons, one economic and one humanitarian. Direct and indirect organizational costs result from mismanaged job stress. In addition, mismanaged job stress causes physical and emotional suffering in the workplace. The National Institute for Occupational Safety and Health (NIOSH) identified job stress and psychological disorders among the top 10 occupational health hazards in the United States and the American Institute of Stress determined that job stress reached epidemic proportions in the by the end of the 1990s. When life expectancy at birth is used as the standard, Americans are passing the stress test. However, although Americans are living longer, they are suffering more. Anxiety disorders affect one in every six people in

©2004 Elsevier Inc.
All rights reserved.

the United States and one in every five employed people in the United Kingdom. In 2001, 35% of Americans reported that jobs were harming their physical and emotional health. Layoffs and downsizing during the 1990s had adverse effects on employee, managerial, and executive stress. Translated literally from the Greek, epidemic means "upon the people." Epidemiology is the basic science and fundamental practice of public health and preventive medicine. Treatment alone is rarely effective in the management of epidemics. Prevention is the preferred mode of epidemic management.

2. CAUSES OF JOB STRESS

Job stress is caused by demands within the work environment or by contextual or nonwork demands. The major categories of job stressors are physical demands, task demands, role demands, and interpersonal demands. In addition, extraorganizational stressors (e.g., marital discord) and transitional factors (e.g., preparing for retirement) can have spillover effects into the workplace, in part through role conflicts. Understanding an individual's stress and strain requires examining the entire life experience (i.e., work plus nonwork).

Physical stressors vary by occupation and working environment. Extreme work conditions, such as in tropical or artic work environments, are a source of physical stress, as are strenuous working conditions, such as in industrial construction or steel making. Business travel is a major physical stressor for those who must engage in this activity.

Task stressors include a wide range of job content and job context factors. Work overload, time pressure, and changing tasks or activities are among the greatest sources of job stress for people in many occupations and organizations. Four other task-related sources of job stress are lack of control, career progress, future job uncertainty, and new technologies. Lack of control and lack of decision latitude are major contributors to job strain, especially when work demands are very high for a particular job. Blocked career progress leads to strain from frustration of natural achievement drive and ambition, whereas future job uncertainty creates strain from the anxiety of job loss. New technologies cause job stress due to the requirement for developing one's abilities and learning new skills.

Role stressors are among the best understood sources of job stress. The two major categories of roles stressors are role conflict and role ambiguity. Role conflict may emerge in a variety of ways. Interrole conflict occurs when the expectations of different roles (e.g., the role of employee and the role of wife) are in conflict with each other. Intrarole conflict occurs when the expectations from different role senders (e.g., one's supervisor and one's colleague) are in conflict with each other. A variation of intrarole conflict occurs when one role sender, such as one's supervisor, sends conflicting expectations. Person-role conflict occurs when one's personal beliefs or values are in conflict with the expectations of one's job. Role ambiguity occurs in several forms. Ambiguity may arise when it is unclear what the consequences might be for failure to meet others' expectations, or it may arise as to exactly what others' expectations really are or what they really mean.

Interpersonal stressors are somewhat broader and more personal than role stressors. Leadership styles, abrasive personalities, organizational politics, poor diversity management, and sexual harassment give rise to interpersonal stress on the job. The two most stressful leadership styles for people at work are rigid, autocratic leadership behaviors and laissez-faire, or very passive, leadership behaviors. Abrasive personalities can be the source of intense interpersonal stress and emotional pain on the job. Organizational politics is a source of job stress, with some research showing that it has a greater adverse impact on women than men. Poor diversity management leads to job stress for the minority worker who feels there are unequal workplace barriers to success. Sexual harassment is a major job stressor, most often for women. In addition to being significant interpersonal stressors, diversity issues and sexual harassment can lead to tens of millions of dollars in legal costs and settlements.

Nonwork stressors, such as marital and family expectations, child-rearing responsibilities, and parental care, create role conflicts for people on the job and have spillover effects into the work environment. Traumatic events outside of the workplace, such as the loss of loved ones or the reality of terrorism and war, also may have spillover effects into the workplace, thus magnifying any job stress for the individual.

3. CONSEQUENCES OF JOB STRESS

The two broad categories of job stress consequences are functional (eustress, or healthy stress) and dysfunctional (strain). The former includes both health and performance benefits. The health benefits of well-managed job stress include enhanced focus during peak work activities and workplace emergencies, as well as

cardiovascular efficiency from aerobic fitness. The performance benefits include enhanced activity level and arousal with their accompanying bursts of physical strength and peak performance. The state of active engagement in meaningful work that produces eustress represents the healthy side of stress consequences. Workers who experience eustress can translate their experience into benefits for the organization, including productivity, organizational agility, and efficiency.

The dysfunctional consequences of job stress are individual and organizational strain. Because job stress triggers a systemic response within individuals and organizations, individual and organizational strain may manifest in a wide variety of ways. Job strain's manifestation is importantly influenced by the particular vulnerability of the individual or weak link in the organization. This is commonly called the Achilles' heel phenomenon or weak organ hypothesis. Some individuals, for example, tend to respond to stress via the respiratory system. They may find that their allergies worsen during periods of high stress. In some organizations, the weak link may be ethical behavior, with more unethical actions occurring during periods of high stress. For both individuals and organizations, it is at the point of the Achilles' heel that many strain symptoms emerge.

Individual strain can take the form of psychological disorders, medical illnesses, and/or behavior problems. Anxiety and depression are two of the more common psychological consequences of strain. Individuals may also experience family problems and sleep disturbances related to work strain. Within the services sector, individuals are particularly prone to burnout, another psychological consequence. Among the medical consequences of strain are heart disease and stroke, cancer, diabetes mellitus, and emerging diseases such as HIV and chronic fatigue syndrome. Behavioral consequences include tobacco and alcohol abuse, eating disorders, and violence.

Because of the exchange relationship between employees and organizations, the consequences of strain for employees translate into costs for the organization. Organizational strain can take the form of direct costs and/or indirect costs. Direct costs include strikes and work stoppages, declining quality and quantity of performance, compensation awards, and increased health care costs. Among the indirect costs of distress are employee dissatisfaction, communication breakdowns, faulty decision making, and diminished quality of work relationships.

4. MODIFIERS OF THE STRESS RESPONSE

The stress response is the psychophysiological pattern that results when an individual is exposed to stressors. The stress response pattern is universal; the intensity, frequency, and consequences of the stress response are not. The ways in which the stress response is exhibited in any individual are moderated by individual and interpersonal factors, including interpersonal communication abilities, personality differences, emotional intelligence, and gender differences. The interaction of these variables with the stress response determines whether the stressor results in positive outcomes (eustress) or negative outcomes (strain). Although many of these moderators are inherent traits, most can also be consciously changed through prevention efforts.

Poor interpersonal communication has been cited as one of the precursors to strain. Therefore, an individual's ability to effectively communicate when faced with stressors can positively influence the outcomes of the stress. Interpersonal communication can be in the form of effective management of difficult work situations, people, or role conflicts. It can also involve effective use of communication forums to gain social support and to positively funnel stress-induced emotions. Effective interpersonal communication skills are in some cases inherent skills related to personality variables, but they can also be acquired through active learning of productive communication processes.

Personality differences also modify the individual's response to stressors and the consequences of the individual's response. Personality affects the person's perception or appraisal of a specific situation as stressful and also affects the person's response to the perceived stressor. Although multiple personality traits influence the stress response, the most often cited is Friedman and Rosenman's type A vs type B behavior patterns. Research has consistently shown that type A individuals exhibit more distress than their type B counterparts. The type A individual's high-intensity work level, incessant need for achievement, inability to relax, and general tendency to lack a balanced life approach result in both a higher number of situations being perceived as stressful and poor coping responses when stressors occur. Another perspective on personality and stress is illustrated by the Myers–Briggs Type Indicator. This popular measurement tool for personality traits depicts a combination of traits that describe how a

person perceives the world, processes information, and makes decisions. The individual's personality type explains why one person who is comfortable with less structure, independent work, and frequent changes may have a more positive response to an organization in flux than an individual who prefers structure, stability, and has a strong need for closure. Two other personality variables linked to the stress response are hardiness and self-reliance. Hardiness is defined as responding to stressful situations with commitment (vs alienation), control (vs powerlessness), and challenge (vs threat). Self-reliance is defined as the ability to form and maintain multiple healthy and mutually supportive interpersonal attachments that are used appropriately to deal with stressors. Some personality traits may be malleable, such as "learned optimism," and thus can be developed in order to help individuals manage stress more effectively.

Emotional intelligence (EI) includes a set of competencies that refer directly to the way an individual deals with emotions, a key determinant of how they also deal with stress. The components of EI include the abilities to be self-aware of emotions, to express them appropriately, to judge sincere vs insincere expressions of emotions, to distinguish and prioritize different emotions, and to assess problems from multiple emotional perspectives. EI components also include the ability to understand complex dual emotions and natural transitions from one emotion, such as fear, to another, such as anger or hostility, and to connect or disconnect from an emotion depending on its usefulness to the situation. EI may naturally occur in some individuals, but it can also be improved through targeted learning.

Gender differences also moderate the stress response. Research suggests that the natural response to stressors differs for women from the traditional fight-or-flight pattern. In addition to the fight-or-flight response, women "tend and befriend" in times of high stress. The moderating effects of tend and befriend may explain differences in psychophysiological consequences of stress for men and women. There are also issues related to the conflict between work and family roles that appear to affect women to a larger extent than men.

5. MEASURES OF JOB STRESS

Valid, reliable job stress measurement requires multiple psychological, environmental, and medical measurements. There are four construct categories for stress instruments: measures of (i) environmental demands and sources of stress; (ii) healthy, normal stress response; (iii) modifiers of the stress response; and (iv) psychological, behavioral, and medical strain. For the purposes of workplace stress measurement, the focus is on four measures that offer a range of construct assessment. For a more clinical assessment of organizational stress, Levinson's classic organizational diagnostic approach may be used.

5.1. Job Stress Survey

Spielberger's Job Stress Survey (JSS) measures job pressure and lack of organizational support as the two primary sources of job stress. The JSS measures the severity and the frequency of 30 different job demands. Examples from the JSS scale items are "excessive paperwork" and "poorly motivated coworkers." Severity and frequency scores are also computed for 10-item job pressure and organizational support subscales that were factor analytically derived from the 30 JSS items. The subscales show adequate internal consistency and reliability.

5.2. NIOSH Generic Job Stress Questionnaire

The Generic Job Stress Questionnaire was developed by NIOSH and contains measures of 13 different job stressors as well as a host of measures of individual distress and modifiers of the stress response. The questionnaire assesses constructs within domains contained in the NIOSH job stress model. Specific stressor, strain, and modifier variable constructs were selected for inclusion in the instrument on the basis of a content analysis of the job stress literature, and the scales selected to measure these constructs were adapted from scales with known reliability and validity. The instrument was designed to be modular in form so that the diagnostician or stress researcher can select individual scales, or the entire instrument can be used.

5.3. Occupational Stress Inventory

The Occupational Stress Inventory (OSI) measures three domains of adjustment: occupational stress, personal strain, and coping resources. Six occupational stress scales measure role overload, role insufficiency, role ambiguity, role boundary, responsibility, and the physical environment. Four personal strain scales

measure vocational strain, psychological and/or emotional problems, interpersonal problems, and physical illnesses. Four coping scales measure recreational activities, self-care, social support, and rational/cognitive coping. Plotting standardized scores on each of the 14 scales produces a "stress profile" for an individual manager or executive, or the scales can be used to develop organizational profiles.

5.4. Pressure Management Indicator

The Pressure Management Indicator (PMI) is a measure developed from the earlier OSI and based on the Cooper and Marshall's stress model. The PMI is used increasingly as a diagnostic instrument in Europe. It is more reliable, more comprehensive, and shorter than the OSI. The PMI measures sources of pressure, such as workload, relationships, work–home balance, and daily hassles; individual differences, such as type A behavior/drive, locus of control, and social support; and job stress effects, such as job satisfaction, security, mental health, and physical symptoms.

6. EARLY WARNING SIGNS OF STRAIN

From the prevention perspective, it is important to recognize the earliest warning signs of strain in order to intervene early in the process. Worker strain can result following prolonged exposure to stressors when paired with poor response patterns. Pains of unknown origin, fatigue, inability to concentrate, and irritability are early warning signs for individuals. In addition, symptoms of depression, increased anger and hostility, increased accidents, or signs of substance abuse can be clues that interventions are warranted. These early warning signs may be accompanied by increased aggressive behavior, disrespect, and withdrawal from relationships.

For organizations, early warning signs that normal work stressors are leading to strain include general patterns such as slight changes in productivity with decreased quantity and/or quality of work, increased absenteeism, decreased commitment to the organization, poor interpersonal work relationships, increased tardiness, and more conflict among workers. Other early signals of distress include general dissatisfaction, low motivation, and low morale. Increased accident rates and machine breakdowns may also constitute early warnings. There may be a general loss of vitality within the organization and an atmosphere of distrust and animosity.

As with any preventive strategy, early detection of strain can help to reduce the long-term negative consequences. Key to early detection and sound prevention is vigilance on the part of the management team.

7. PREVENTIVE STRESS MANAGEMENT

Prevention is the best public health strategy for any disease epidemic. Because job stress is a health epidemic, prevention holds the best hope for addressing this epidemic. The theory of preventive stress management translates the public health notions of prevention into an organizational context and overlays them on a stress process model as shown in Fig. 1. Stress is one of several chronic organizational health problems—with others being workplace violence, sexual harassment, and suicide—for which prevention is appropriate.

Preventive stress management is an organizational philosophy and set of principles that employs specific

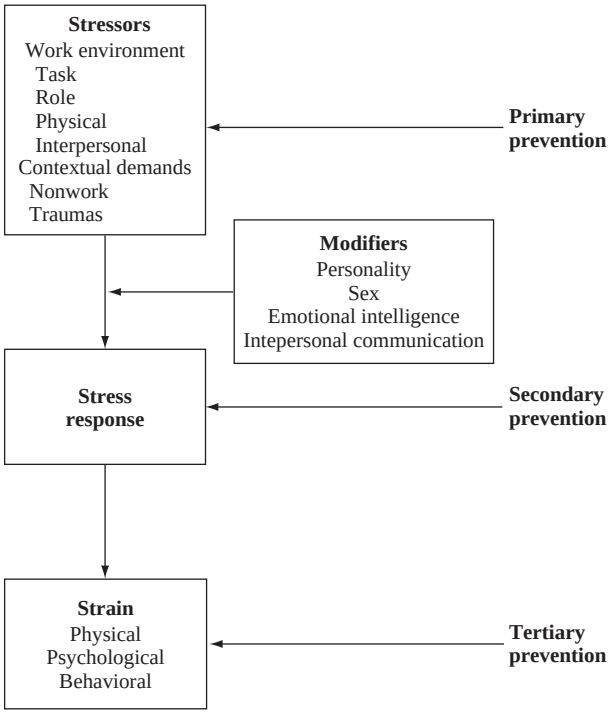


FIGURE 1 Preventive management model for job stress.

methods for promoting individual and organizational health while preventing individual and organizational distress. This philosophy is based on the following five guiding principles that motivate the practice of preventive stress management and provide a framework for healthy organizations and healthy leaders:

Principle 1: Individual and organizational health are interdependent.
Principle 2: Leaders have a responsibility for individual and organizational health.
Principle 3: Individual and organizational distress are not inevitable.
Principle 4: Each individual and organization reacts uniquely to stress.
Principle 5: Organizations are ever-changing, dynamic entities.

Chronic diseases do not arise suddenly; instead, they develop gradually through a progression of disease stages, a "natural life history." This is true for chronic individual disorders, such as heart disease, as well as for chronic organizational problems, such as workplace violence. The natural history of most diseases is one of evolution from a stage of susceptibility to a stage of early disease and finally to a stage of advanced or disabling disease. At the stage of susceptibility, the individual is healthy but is exposed to certain risk factors or disease precursors. For example, individuals who choose a sedentary life or who choose to smoke cigarettes are at the stage of susceptibility for coronary artery disease as well as several other diseases. When these and other risk factors lead to the development of arteriosclerosis or hardening of the arteries to the heart, the individual is at the stage of early disease or preclinical disease. In other words, the person's body has responded to the disease precursors, but there are few, if any, symptoms. As the disease advances, it becomes symptomatic or clinical disease. Angina pectoris (heart pains) and heart attacks are advanced manifestations of coronary artery disease.

Preventive stress management is rooted in the public health notions of prevention, which were first used in preventive medicine. The term public health encompasses a broad array of health protection activities inspired by the practice of viewing illnesses within a social context. The dominant diagnostic model in public health involves the interaction between a host (the individual), an agent (health-damaging organism or substance), and the environment. One of the fundamental concepts of preventive medicine is that there is an opportunity for preventive intervention at each stage in the life history of a disease, as noted previously and depicted in Fig. 1. These interventions are aimed at slowing, stopping, or reversing the progression of disease.

As Fig. 1 shows, there are three stages of prevention strategies: primary prevention, secondary prevention, and tertiary prevention. Primary prevention aims to modify and manage the job stressors and other demands in the work environment. Secondary prevention aims to modify and manage the individual's response to these job stressors and other demands. Tertiary prevention aims to help and provide aid to those who are experiencing behavioral, psychological, or medical strain symptoms. From a public health perspective, primary prevention is always the preferred point of intervention. For job stress and workplace health, this implies that job redesign efforts and other interventions that alter, modify, or eliminate stressful work conditions are the preferred category of preventive stress management interventions.

Primary prevention is the protection of health directed at the stage of susceptibility and aims to eliminate or reduce the impact of risk factors; it is intervention before the onset of problems or disorders. Primary prevention may be either organizational or individual. Organizational strategies include the following:

Job redesign
Participative management
Flexible work schedules
Career development
Design of physical settings

Individual strategies include the following:

Learned optimism
Planning and time management
Social supports and secure attachments
Maintaining a work–home balance
Changing type A behavior pattern

It is important to recognize that some primary prevention techniques directly alter stressors. Job redesign, for example, may incorporate more control into a job, or more decision latitude, thereby reducing the stress from lack of control. On the other hand, some primary prevention methods alter the individual's perception of the stressor. Social support, for example, can help individuals reframe stressors as challenges because they believe they possess the resources to manage a stressor. In both cases, pri-

mary prevention is the preferred first defense against stressors.

Secondary prevention aims at the early detection of problems or disorders and prompt, early interventions to correct departures from health. Secondary prevention may also be either organizational or individual. Organizational strategies aim at improving relationships in the work environment. Organizational strategies include the following:

Goal setting
Role analysis
Team building
Social supports and secure attachments
Diversity training and programs

Individual strategies include the following:

Relaxation training
Prayer, spirituality, and faith
Emotional outlets
Physical fitness and exercise
Nutrition

Secondary prevention efforts focus on the development of strengths for both individuals and organizations. Team-building efforts, for example, can lead members to understand each other better, manage conflict more effectively, and be more productive. Individual strategies, such as relaxation training, meditation, and exercise, make the body and mind more resilient and more capable of dealing with stress.

Tertiary prevention is prevention of more serious problems, disorders, or death. It is directed at expediting or improving treatment for symptomatic or advanced disease and aims to alleviate discomfort and restore effective functioning. Tertiary prevention is primarily focused at the individual level rather than the organizational level, as is done in treating mental health referrals. Tertiary prevention also includes psychological counseling and medical or surgical treatment. Individual strategies include the following:

Individual psychotherapy
Behavior therapy
Medical care and treatment
Traumatic event debriefings
Physical therapy

Tertiary prevention is thus focused on procuring care from qualified professionals who can help individuals heal. The goal of the preventive stress management approach is to utilize sufficient primary and secondary prevention such that individuals seldom, if ever, need to turn to tertiary prevention for relief.

8. IMPLICATIONS FOR EXECUTIVES AND LEADERS

Job and work stress, as a component of the contemporary organization, do not show signs of decreasing in frequency, intensity, complexity, or outcomes. Effective management of organization and worker stress by executives and leaders begins with the effective management of their own stress. Learning preventive stress management can benefit the individual leader and positively affect stress management throughout the organization. As the workforce continues to change with increased globalization, diversity, and age variation, effective leaders must learn better ways to deal with the inevitable stressors of the workplace. Future leaders may very well set themselves apart by developing the stress prevention skills needed to move the response to stress in a positive direction so that individuals and organizations can reap the benefits of eustress and avoid the pitfalls of strain.

The key tasks for managers and executives are clear. Within the preventive stress management framework, the goal is to increase eustress—the healthy, positive consequence of stress—and to prevent or resolve distress. The three-tiered model presented in Fig. 1 can be used as a road map for these prevention efforts. Primary prevention represents the most efficacious set of techniques because it consists of modifying the stressor or the individual's response to the stressor. Secondary prevention efforts, focused on increasing individual and organizational resilience, should supplement primary prevention methods. Finally, tertiary prevention methods are available for targeting symptoms and healing the wounded within organizations.

Increasing eustress within organizations presents a challenge to managers and to researchers, who are beginning to identify the aspects of work that many employees find pleasurable and motivating. The feeling of eustress has been described as being in the "flow," similar to a "runner's high" in which time suspends and the person is fully absorbed in the joy of work. It is a state of active engagement, zest, and achievement. Eustress is characterized by feelings of hope,

commitment, and meaningfulness of work. Managers and executives must explore ways to design jobs that tap this positive experience for employees. Managers should also find opportunities to improve policies, procedures, and the physical work environment in an effort to enhance employees' exposure to work that they find engaging and to eliminate potential barriers to eustress.

Executives are vulnerable to the same medical, psychological, and behavioral consequences that affect any workers. They are at increased risk for isolation and loneliness, anxiety, burnout, workaholism, and depression. Their positions in organizations place them as role models whom others look to for guidance regarding effective stress management. Because of their positions, they can model effective prevention efforts that others can follow, thus benefiting the entire organization. By taking a holistic view of health and developing its multiple dimensions (physical, psychological, spiritual, and ethical), executives can lead the way in tapping into the benefits of eustress and possessing the strength and security to effectively manage distress.

Because job stress leads to a systemic response both within the organization and within the individual, executives and workers alike can benefit from strengthening themselves through physical fitness, psychological well-being, spiritual development, and strong character. Physical fitness includes cardiovascular well-being, muscular flexibility, and appropriate muscular strength. Psychological well-being centers on the ability to deal constructively with reality, the capacity to adapt to change, relative freedom from tension or anxiety-related symptoms, and the capacity to love. Spiritual development should not be confused with religious affiliation and is characterized by hope, a sense of purpose, and healthy relationships. Finally, strong character is achieved through ethical behavior and positive personal integrity.

In addition, executives can be proactive in creating healthy work environments in which job stress is optimized, rather than either maximized or minimized. High-performance, healthy work environments are ones characterized by a sense of challenge coupled with a sense of psychological safety and security. Healthy job stress leads to a work environment in which people may produce, serve, grow, and be valued.

Future areas of study in work and job stress include exploration of the individual differences in the stress response and how to incorporate the positive traits into effective stress prevention. Training in skills shown to increase the positive outcomes of stress and reduce strain needs to be incorporated into educational programs at an early stage and throughout an individual's work life. Finally, shifting the perception of job stress to a neutral construct is both necessary and inevitable. The human response to stress, not the stress itself, determines whether negative or positive outcomes prevail.

See Also the Following Articles

Environmental Stress ■ Homosexuality ■ Job Stress and Burnout ■ Stress

Further Reading

Bar-On, R., & Parker, J. D. A. (2000). *Handbook of emotional intelligence: Theory, research, assessment and application at home, school and in the workplace*. New York: Jossey-Bass.

Cooper, C. L. (1996). *Handbook of stress, medicine and health*. Boca Raton, FL: CRC Press.

Hurrell, J. J., Jr., Nelson, D. L., & Simmons, B. L. (1998). Measuring job stressors and strains: Where we have been, where we are, and where we need to go. *Journal of Occupational Health Psychology*, *3*, 368–389.

Levinson, H. (2002). *Organizational assessment: A manual*. Washington, DC: American Psychological Association.

Macik-Frey, M., Quick, J. C., & Quick, J. D. (In press). Interpersonal communication: The key to unlocking social support in preventive stress management. In C. L. Cooper (Ed.), *Handbook of stress, medicine and health* (Rev. ed.). Boca Raton, FL: CRC Press.

Perrewé, P. L., Ferris, G. R., Frink, D. D., & Anthony, W. P. (2000). Political skill: An antidote for workplace stressors. *Academy of Management Executive*, *14*(3), 115–123.

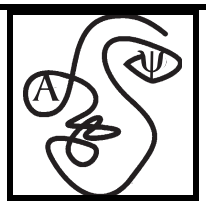
Quick, J. C., Cooper, C. L., Nelson, D. L., Quick, J. D., & Gavin, J. H. (2003). Stress, health, and well-being at work. In J. Greenberg (Ed.), *Organizational behavior: The state of the science* (pp. 53–89). Mahweh, NJ: Laurence Erlbaum Associates.

Quick, J. C., Cooper, C. L., Quick, J. D., & Gavin, J. H. (2002). *The Financial Times guide to executive health: Building your strengths, managing your risks*. London: Financial Times/Prentice Hall.

Quick, J. C., Quick, J. D., Nelson, D. L., & Hurrell, J. J., Jr. (1997). *Preventive stress management in organizations*. Washington, DC: American Psychological Association. (Original work published 1984 by J. C. Quick & J. D. Quick).

Salovey, P., Rothman, A. J., Detweiler, J. B., & Steward, W. T. (2000). Emotional states and physical health. *American Psychologist*, *55*, 110–121.

Taylor, S., Klein, L., Lewis, B., Gurung, R., Grunewald, T., & Updegraff, J. (2000). Biobehavioral responses to stress in females: Tend and befriend, not fight or flight. *Psychological Review*, *107*, 411–429.



Job Stress and Burnout

Philip Dewe
Birkbeck University of London, London, United Kingdom

1. Introduction
2. Defining Job Stress
3. Models of Job Stress
4. Causes of Job Stress
5. Coping with Job Stress
6. Burnout
7. Other Job Stress Responses
8. Job Stress Interventions
Further Reading

GLOSSARY

appraisal The processes an individual goes through in determining the significance and meaning of a stressor (primary appraisal) and what resources are available to cope (secondary appraisal).

strain Responses that indicate that the individual is under stress.

stress management programs Interventions at the level of the individual and organization aimed at reducing stress and allowing the individual to fully contribute to the goals of the organization.

stressors Situations or properties of a situation that have the potential to cause stress.

Job stress has traditionally been defined in terms of a stimulus, response, or an interaction between the two. These approaches to defining job stress have allowed researchers to identify an array of different causes of job stress (stressors) and relate them to different responses (strain). Contemporary definitions of stress view it as relational, resulting from the transaction between the individual and the environment. This approach has provided an opportunity to consider those processes that link the individual to the environment. The shift in emphasis away from the different components of the stress process to the process itself has raised a number of issues in terms of how measurement practices should develop and how researchers fulfill their obligations to those whose working lives they study.

1. INTRODUCTION

The growing awareness of the impact of job stress is best illustrated by the sheer increase in the volume of job stress research in recent years. Job stress is a condition familiar to many. In some form or another, it influences all working lives, leaving those who experience job stress emotionally drained and frequently more vulnerable to other illnesses and disease. It is not just individuals who experience the impact of job stress. Organizations, families, communities, and society all bear the costs of job stress to such an extent that job stress is frequently viewed as an inevitable feature of 21st-century living. However, when sifting through all the information on job stress it is important to maintain a sense of perspective. It is necessary to guard against feeling that the problem of job stress has been overexaggerated, that it is not worth worrying about, or that it is someone else's responsibility. The

©2004 Elsevier Inc.
All rights reserved.

issue is not whether there is too much or too little stress in our working lives but what actually causes job stress, what can be done to cope with it, and what needs to be done to provide a working environment within which individuals can maximize their potential and enhance their well-being. This is the responsibility that job stress researchers have to those whose working lives they study.

2. DEFINING JOB STRESS

The difficulty when talking about job stress is that it means different things to different people, with the result that the term stress has taken on a variety of meanings. Traditionally, stress research has been influenced and guided by definitions that focus on stress as a stimulus, a response, or on the interaction between the two. These three approaches to defining stress frequently reflect the discipline or orientation of researchers. Stress as a stimulus is defined in terms of those events or properties of events that place demands on an individual over and above what is normal. Response definitions of stress are concerned with identifying particular responses or patterns of responses that indicate that the individual is confronted with a demanding event. Stimulus and response definitions of stress have provided researchers with an opportunity to identify both a range of events that have demanding properties and responses that indicate that the individual is under stress. However, such definitions have been criticized because each reflects just one aspect of the stress process; defining stress in this way fails to provide any understanding of the nature of the stress process. They have also been criticized because they fail to take into account individual differences and therefore cannot account for the fact that what may be stressful for one individual may not be stressful for another.

Interactional definitions of stress tried to overcome these difficulties by describing stress in terms of the interaction between the stimulus and the response. This definition is one in which a relationship, usually correlational, is hypothesized between the stimulus and the response. Cooper *et al*. argue that defining stress like this, where the focus is simply on the interaction between the stimulus and the response, fails to provide any understanding of the complexity of the interaction or of the nature of the processes that may be involved. As a consequence of the difficulties associated with stimulus, response, and interactional definitions of stress, more contemporary definitions adopt a definitional approach described by Lazarus in which stress is viewed as relational in nature involving some sort of a transaction between the individual and the environment. Transactional definitions of stress imply that stress is neither solely in the individual nor in the environment but in the transaction between the two. It is the transaction that links the individual to the environment. Furthermore, it is by thinking of stress in transactional terms that draws attention to what the nature of the transaction may be. In this way, defining stress in transactional terms achieves what earlier definitions failed to do: It focuses attention on those processes that link the individual and the environment and that are at the heart of the transaction.

Lazarus describes these processes as primary and secondary appraisal. Primary appraisal is when the individual evaluates the importance of the event in terms of harm, threat, or challenge to well-being. In short, it is the realization by the individual that something important is at stake. Secondary appraisal is when the individual considers what can be done. It involves an evaluation of those resources that are available to cope with the stressful event. Identifying the processes that are at the heart of a stressful transaction in terms of how an event is appraised and how it can be coped with means that contemporary definitions of stress are more concerned with capturing its transactional qualities than focusing, as was once the tradition, on individual components to that transaction. As a result, stress is best defined in terms of those encounters in which the demands from the encounters are appraised as taxing the individual's resources, thus threatening well-being.

3. MODELS OF JOB STRESS

Models of job stress have, quite naturally, closely followed how stress has been defined. Early models of job stress followed an interactional framework. That is, they were interested in the interaction between the cause of job stress (stimulus) and its impact (response). As outlined, the interactional framework suggests a relationship between the perceived presence of demanding work conditions (stressors) and different stress responses (strain). The relationship, usually correlational, postulated that the more demanding the stressor, the greater the probability it would result in strain. As job stress researchers began to explore different relationships between stressors

and strains, they soon realized that the stressor–strain relationship may be influenced by other variables, such as age, gender, organizational level, and individual differences. Therefore, a second stage in the development of job stress models began. This stage is described as the moderator stage and was concerned with identifying organizational and individual variables that moderate (influence) the stressor–strain relationship.

The interactional and moderator models, although important in providing information on the causes and consequences of job stress and on variables that influenced the relationship between the two, were limited in their ability to explain the stress process. So began in the development of job stress models the transactional stage. Transactional models, like transactional definitions of job stress, were concerned with the sequence of events and the processes that linked the stressor to the strain. McGrath had for some time urged researchers to approach the investigation of job stress by developing models of job stress that mirrored the sequencing of events from stressor to strain. To capture this sequence, many models of job stress adopted the concept of fit. The idea of fit was coupled with the idea of equilibrium or balance. In general, where there was a misfit or imbalance between the person and the environment, such as in the case of a person failing to cope with the demands of a job, a state of disequilibrium would exist and this state would be associated with strain. Implicit in the concept of misfit is the notion of an individual's ability to manage or deal with a demanding job event.

As job stress researchers began to explore the nature of this misfit, models of job stress began to outline more explicitly what misfit may involve. Job stress models began to build into their frameworks different transactional qualities, such as how stressors are appraised, how individuals cope, and organizational and individual resources available to individuals to manage the demanding encounter. What distinguishes the transactional model from the earlier interactional and moderator models is its focus on process. It requires researchers to identify the processes that link the individual to the environment and to consider the sequencing of those processes and their role over time. However, although job stress researchers have long accepted at the theoretical level the importance of identifying process elements in job stress models, empirical work is predominately still being carried out using a interactional–moderator perspective. Nevertheless, job stress research has contributed much to our understanding of its causes and consequences, not at times without controversy and intense debate.

4. CAUSES OF JOB STRESS

The causes of job stress are many and varied. Perhaps the easiest way to develop a comprehensive picture of what causes job stress is to start with the early work of Kahn and colleagues and two causes of work stress: role conflict and role ambiguity. These two work role demands along with role overload have become the most frequently researched causes of job stress. In general, role conflict refers to job situations in which individuals are confronted with conflicting demands while performing their organizational role so that by complying with one of the demands it is very difficult to comply with the other(s). Role conflict frequently arose from having to work under policies that conflicted with one another, being responsible for tasks over which there was very little control, and having to do things that were likely to be accepted by one person but not accepted by another. Perceived role conflict led many to experience dissatisfaction and strain. Similar findings have been found for role ambiguity. In the case of role ambiguity, the significant issue was a lack of information. Individuals experiencing role ambiguity more often than not found that they lacked the relevant information to perform their work roles, determine how their performance was being judged, and consider whether the decisions they made were the right ones. Role ambiguity was consistently associated with levels of psychological strain.

A third form of difficulty associated with performing one's role in the organization was role overload. Role overload was first measured in terms of the demands that arose from the volume of work that had to be done. Frequently associated with time pressures, this type of role overload became known as quantitative overload and was measured in terms of the amount of work to be done in the available time, the idea that all tasks have a sense of urgency, not being able to do the work because of constant interruptions, and the belief that no matter how much work was done there was still as much to be done at the end of the day as there was at the beginning. As organizational roles became more complex, role overload took on a second dimension. This type of role overload was described in terms of qualitative role overload and referred to demands on individuals required to ensure that their skills matched the increasing requirements of the job. This form of

role overload was concerned with the need to keep up with the technical requirements of the job, having to spend increasingly more time learning new policies and procedures, keeping up with constant job changes, and sensing that at times there was an imbalance between current skills and required skills. Both these forms of role overload had negative consequences in terms of well-being and satisfaction on those who perceived them to be present in their jobs.

The importance of discussing these types of role demands in such detail is that even though they were first identified in the 1960s, they are—despite the significant social, economic, political, and organizational changes that have occurred since that time—still by far the most frequently measured causes of job stress. Other causes of job stress have been identified, and it is worth exploring these as well. In addition to organizational roles, in 1976 Cooper and Marshall identified a range of different causes of job stress, including extrinsic aspects of the job, such as working conditions and job design; structure and climate issues, such as levels of participation, supports, controls, and rules and regulations; career development issues, such as promotion, job security, and goal achievement; and relationship issues involving, for example, subordinates, colleagues, and the boss, and difficulties in delegating and dealing with sensitive interpersonal issues.

Cooper and Marshall also included "extraorganizational sources of stress" in their list of causes of job stress. These sources relate to the home–work interface and include balancing family and work responsibilities; family problems; financial and other life difficulties; and the interference of work with nonwork activities, such as leisure, hobbies, community life, and social activities. Other sources of job stress include the level of local and international travel, working abroad, life-threatening events and safety at work, redundancy, bullying, harassment, social responsibility, and ethics. Beehr also provided a similar list of the causes of job stress.

Although a wide range of events, usually defined as stressors, can be identified as having the potential for causing job stress, there has been considerable debate about how best to measure such stressors. The common approach has been to use self-report measures in which individuals generally indicate on a 5-point scale whether they agree or disagree as to the presence of different stressors in their job. The major criticism of this approach is that simply agreeing or disagreeing that a stressor is present in a job does not actually indicate how demanding that stressor may be or its significance in terms of well-being or complexity. As a consequence, a number of refinements to self-report measures have been suggested. First, there has been discussion as to whether stressors should be measured globally, allowing comparisons to be made across different occupational groups, or tailored to meet the requirements of specific occupations so that the unique stressors for that occupation can be identified. This discussion, particularly in relation to global measures, often spills over into what events should be measured. That is, have job stress researchers, by focusing too much on organizational role stressors, overemphasized the presence of these work role events and ignored the significance of other events, particularly those that reflect the enormous change that has occurred in society and organizations?

Second, when measuring stressors should more attention be paid to the characteristics of the stressors? This would include measuring the frequency, duration, and intensity of the event; how demanding it is; and the way in which different events may result in the accumulation of stress. Then, thinking of stress in transactional terms, there is a need to consider the difference between the presence of an event and its significance and meaning. Here, the emphasis shifts to determining how an event is interpreted or appraised. This means developing measures that capture the meanings individuals give to the event. In this way, the focus would be on whether the event is appraised in terms of, for example, harm/loss and threat or challenge. Focusing on the meaning and salience of an event may be a more significant predictor of job stress than the more traditional measurement approach that simply captures whether the event is present in the job. Of course, focusing on the meanings that individuals give to events raises the issue of subjective versus objective measurement and the fact that objective measures are also an important component contributing to our understanding of the causes of job stress.

Debating how to refine self-report measures of stressors is one way of capturing the complexity and character of different stressors. However, as Cooper *et al.* suggest, there are other more qualitative ways of measurement that should not be overlooked, including the use of critical incident techniques, daily diaries, and structured interviews. These measurement strategies allow stressors to emerge from the narrative allowing individuals to identify the significant events in their working lives. Whatever measurement method is selected, it is clear that identifying stressors is a crucial

part of furthering our understanding of the stress process and in considering the development of different interventions strategies.

5. COPING WITH JOB STRESS

When stress is defined in transactional terms, coping is fundamental to our understanding of the stress process. Coping is part of what Lazarus describes as secondary appraisal. Secondary appraisal is concerned with "what can be done" to deal with the demanding situation and involves identifying and evaluating the available coping resources. Secondary appraisal therefore involves what a person actually thinks and does; thus, as Dewe *et al*. suggest, coping can be defined as "the cognitions and behaviors, adopted by the individual following the recognition of a stressful encounter, that are in some way designed to deal with that encounter or its consequences." From their review of the coping literature, Dewe *et al*. suggest that it is possible to identify three themes associated with defining coping in this way: Coping is (i) relational in the sense that it reflects the relationship between the individual and the environment, (ii) process focused, and (iii) integrative in that it is part of the appraisal process that links the different dimensions to the stress process.

The most common approach to researching coping in relation to job stress is what Cox describes as taxonomic, where researchers describe and categorize different coping actions that appear broadly applicable to work situations. By far the most popular schema for classifying coping strategies distinguishes between what Lazarus and Folkman describe as problem-focused and emotion-focused strategies. Problem-focused strategies are primarily concerned with problem solving or altering the source of the problem, whereas emotion-focused strategies are aimed at reducing the emotional discomfort. This is not the only schema for classifying coping strategies. In general, however, most classification attempts describe variations on the problem- and emotion-focused themes. Nevertheless, O'Brien and DeLongis suggest that any classification of coping strategies should include what they describe as relationship-focused coping. They argue that it is important to recognize that coping requires, in addition to problem solving and emotional management, relationship maintenance, and that coping classifications should be expanded to include this type of coping. From their more detailed examination of coping functions, Ferguson and Cox suggest that there is a degree of consensus among researchers that there may be four qualitatively distinct coping types: emotional regulation (dealing with the emotional consequences of the problem), approach behaviors (aimed at directly confronting the problem), reappraisal (strategies for readjusting and reinterpreting the meaning of the event), and avoidance behaviors (ignoring the existence of the event).

Classifying coping strategies into different categories such as problem- and emotion-focused should not be regarded as indicating that one form of coping is better than another. Placing different strategies into one or another category is not the same as understanding how they work or the manner in which they are used. Researchers have come to realize that, when considered in the context of a particular stressful encounter, what function a coping strategy may be performing is not as easy as to define as independently describing the focus of separately derived categories. Dewe noted that researchers have commented on the fact that the same coping strategy may play a number of different functions depending on the particular encounter. Similarly, the sequencing, patterning, and combined use of different coping strategies all contribute to the difficulties of determining whether a coping strategy is problem- or emotion-focused.

Despite these difficulties, coping research has expanded to explore the influence of personality and gender on the use of different coping strategies, the categories of coping strategies associated with different work stressors, the relationship between different categories of coping and different outcomes, the role of different coping types over the longer term, the nature of the organizational culture and climate and its impact on the use of different coping strategies, and the role of coping as a mediator or moderator in the stress process. Many studies have, somewhat implicitly, drawn attention to the question of coping effectiveness. Regarding the issue of how to determine coping effectiveness, the following question needs to be answered: Effectiveness for whom and at what cost? This has led researchers, when faced with the question of how to establish coping effectiveness, to identify a range of issues that need to be considered, including the need to distinguish between coping and ordinary everyday adapting; short-term versus long-term effectiveness; the difference between effort and effect; the individual's competency to use different coping strategies; the goals an individual may be trying to achieve; the nature of the encounter, particularly in terms of the levels of control an individual may have; and the goodness of fit between how an event is appraised and the coping strategy used.

In addition to the work on identifying specific coping strategies that individuals may use, classifying them, and evaluating their use, research has also focused on "capacity techniques" for coping with job stress. These strategies are not so much designed to deal specifically with a stressful encounter as to develop a person's capacity to deal with a stressful encounter should it occur. This is often done in a number of ways—exercise, relaxation, meditation, biofeedback, and a philosophy of life—but the end result is to provide an inner sense of energy and well-being and a greater sensitivity to the signals our bodies may send us and thus a greater capacity for dealing with and building resistance to stressful encounters. Moreover, these techniques are frequently discussed in terms of lifestyle changes, how they can play a supportive role in conjunction with more specific coping strategies and so become part of a repertoire of coping activities, how they depend to a considerable degree on individual commitment and preference, and how each technique should be reviewed occasionally so that the intended results continue to be positive.

Many of the issues discussed previously, in terms of both classifying coping strategies and determining their effectiveness, have led researchers to consider how best coping strategies should be measured. Again, researchers have considered changes to self-report coping instruments, including considering the type of instructions given so as to differentiate between what individuals may generally use and what they specifically use in a particular situation, the way in which coping strategies are worded so as to avoid ambiguity in meaning, and whether scoring coping strategies in terms of the frequency with which they are used best captures the nature of the coping process. The general consensus from the debate on coping measurement is that identifying circumstances that perhaps best suit the use of coping checklists will provide researchers with the opportunity to bring other more qualitative techniques to the study of coping so that our understanding can continue to develop.

6. BURNOUT

Cooper *et al*. describe burnout as "a special form of strain." These authors point to the concept of burnout as reflecting a state of psychological strain initially associated with those working in the human service professions. Since the early studies on burnout, researchers have extended their investigations of the phenomenon to working life in general, considering in detail the consequences of burnout for individuals and their organizations. In the early 1980s, Maslach provided a description of burnout involving three major elements: emotional exhaustion described in terms of not having the emotional energy to sufficiently manage the encounter; depersonalization, in which individuals simply become seen as objects and are treated in a detached way; and a lack of personal accomplishment, in which the tendency is to devalue performance in negative ways. Since this three-dimensional view of burnout was first proposed by Maslach, much discussion has centered on whether emotional exhaustion is the essential feature of burnout, with the roles played by the other two dimensions being disputed.

Cordes *et al*. describe burnout as a developmental process. These authors go on to describe burnout as a gradual eroding process and note that by emphasizing the process of burnout, researchers and organizations are provided with a mechanism for understanding what to look for and the types of interventions that may be necessary. Their work supports a process in which the onset of burnout is marked by emotional exhaustion. Depersonalization follows, as Ashforth and Lee suggest, because it is a "means (albeit futile) of staunching the flow of emotional energy, of coping with growing exhaustion." The issue of whether, as depersonalization occurs, the individual begins to sense a loss of accomplishment and hence a degrading of achievements is less clear. Cordes *et al*. note that one possible reason for this is that a lack of personal accomplishment may also be explained in terms of a range of constraining organizational factors and hence depersonalization may develop somewhat independently of the other two dimensions.

Correlates of burnout are many and varied, and researchers have explored these at a number of levels, including the individual level (e.g., gender, age, commitment, and individual differences), the job level (e.g., work role demands, client relationships, and autonomy), and the organizational level (e.g., organizational culture, management style, and communications). In general, it is clear that a range of individual, job, and organizational level factors influence the experience of burnout. However, it is also clear that more work is still needed to understand where in the process these different factors have their most significant effect, how far their effect can be generalized, and what this means in terms of the development of intervention strategies.

7. OTHER JOB STRESS RESPONSES

The fact that "being under stress" has become such a common phrase has meant that an array of responses (strains) have been measured as indicating that an individual is stressed. The general approach to outlining job-related stress or strain is to divide the different strains into three categories: physiological, psychological, and behavioral. Reviews of physiological indicators used to measure strain would include measures that capture cardiac, biochemical, and gastrointestinal activity. Under the heading of "poor physical health" associated with job stressors, Kahn and Byosiere list increased pulse rate, high blood pressure, high cholesterol levels, ulcers, and cardiovascular heart disease. Physiological indicators may also include, as Jex and Beehr suggest, behaviors that are closely associated with disease, such as smoking and drinking. The use of physiological measures presents researchers with a range of measurement challenges. However, evidence suggests that the demands of the job can produce physiological reactions and so this approach to measurement offers what Cooper *et al*. describe as important causal pathways with considerable explanatory potential.

Not surprisingly, as Jex and Beehr note, psychological strains are by far the most common type of strain measured. To give an idea of the range of psychological strains measured, these authors list strains from recent studies such as depression, alienation, boredom, frustration, anxiety, hostility, and dissatisfaction. All are correlates of job-related stressors. Similarly, Kahn and Byosiere, in their table of "psychological responses to work stress," identify 43 different strains, including in addition to those noted previously, confusion, irritation, resentment, tension, and emotional arousal. Kahn and Byosiere conclude that the psychological effects of job stressors are "real, painful, and costly." Moreover, they extend to health behavior, job performance, and the enactment of other life roles.

Jex and Beehr note that behavioral strains are the least studied. They discuss two broad categories of behavioral strain: organizationally relevant behaviors (e.g., job performance, absenteeism, and turnover) and individually relevant behaviors (e.g., smoking, drinking, and substance use). Similarly, Kahn and Byosiere identify degradation/disruption of the work role, aggressive behavior at work, flight from the job, degradation/disruption of other life roles, and self-damaging behaviors. Both sets of authors agree that behavioral effects both on and off the job are extremely costly to both the individual and the organization and warrant more research. Such behaviors are the result of a complex process, and researchers need to try to develop an understanding of the process and the role that job stressors play in it.

Although different emotional responses are included among the list of psychological strains, there is growing acceptance that more attention should be paid to workplace emotions. If, as Lazarus suggests, stress is essentially about emotional reactions, then exploring the specific role of emotions in the stress process may help to provide a richer understanding of what happens to a person in a stressful encounter. Lazarus explains that it is the way an event is appraised that gives that event its emotional quality. Taking this one step further, Lazarus suggests that different emotions are associated with particular appraisals, thus providing a causal pathway for investigation. Lazarus identifies 15 different emotions and categorizes them as negative (e.g., anger, shame, envy, and guilt), positive (e.g., pride and relief), and mixed (e.g., hope and gratitude).

8. JOB STRESS INTERVENTIONS

The most common approach to describing stress management programs is to divide them into three categories depending on the level at which the intervention takes place. The first category describes primary interventions aimed at reducing job stressors, including, as Newman and Beehr suggest, strategies for changing organizational characteristics (e.g., organizational restructure and redesigning organizational processes), changing role characteristics (e.g., redefinition of roles, increasing participation in decision making, and reducing ambiguity), and changing task characteristics (e.g., better match of individual abilities with job requirements, provide skill training, and job redesign).

The second category describes secondary interventions. The focus here is on providing stress management training. Secondary interventions are probably the most common job stress interventions and involve a range of techniques, strategies, and skills aimed at equipping individuals to better manage the demands of the job. Newman and Beehr identify four categories of intervention: those aimed at psychological conditions (e.g., planning ahead and realistically assessing one's expectations), those aimed at physical or physiological

conditions (e.g., diet, exercise, and sleep), those aimed at changing one's behavior (e.g., learning relaxation strategies, taking time off, and developing supportive networks), and those aimed at changing one's work environment (e.g., changing to a less demanding job).

The third category describes tertiary interventions, and here the provision is in terms of employee assistance programs (EAPs). Defining EAPs is somewhat problematic because of the array of perceptions and the range of programs. The definition developed by Berridge and Cooper reflects this need to be comprehensive. They define EAPs as

> A programmatic intervention associated with the work context, usually at the level of the individual employee, using behavioral science knowledge and methods for control of certain work-related problems that adversely affect job performance, with the objective of enabling the individual to return to making her or his full job contribution and to attaining full functioning in personal life.

The components of an EAP should reflect a preferred model of practice that reflects the needs of the employees, the goals of the organization, and the underlying principle that for both humanitarian and economic reasons rehabilitation is better than replacement. EAPs are generally operated in two modes: in-house provision of the program and the counseling service or externally located EAP and counseling service. Referrals can be through constructive discussion, in which typically the supervisor or manager makes the referral, or through self-referral, in which the employee initiates proceedings. Most EAPs are developed around a systematic set of procedures designed to identify and respond to personal issues that interfere with work performance. These procedures include a needs analysis to determine the nature of the problem; the development of a program of counseling, advise, and assistance depending on the nature of the problem; a procedure for referrals; a protocol outlining ethical issues, including informed consent, confidentiality, and anonymity; follow-up procedures; and a program of evaluation.

Although EAPs have become an established part of many organizational stress management procedures, less attention has been given to evaluating their impact. A number of reasons have been given, including methodological difficulties in relation to collecting and analyzing what is confidential data, establishing appropriate criteria against which success can be gauged, and establishing what is best practice with regard to carrying out evaluative procedures. Although both qualitative and quantitative criteria have been suggested as criteria for evaluation, the reported findings are mixed. Because of the variety of different programs, it may be best to focus on whether the goals of specific programs have been successful rather than attempt to identify universal criteria that cover all programs.

One of the issues regarding stress management programs is whether many of the programs assume that work environments have the same impact on everyone and therefore many of these programs fall short because they offer only partial solutions, place all responsibility for change on the individual, or fail to take into account the wider organizational structure and context within which behavior takes place. This debate has widened to include discussion about the role that theory should play in the design and development of stress management programs. The thrust of these arguments has been that by viewing stress in transactional terms, stress management programs need to acknowledge that stress does not reside solely in the individual or solely in the environment but in the transaction between the two. In this way, the focus for the design of a stress management program would shift to the relational nature of stress and the need to consider both individual and environmental issues rather than simply placing the responsibility on one. The continuing debate reflects a wider concern held by those researching job stress; that is, what needs to be done in terms of measurement practices and theory development so that we can contribute significantly to those whose working lives we research.

See Also the Following Articles

Environmental Stress ■ Homosexuality ■ Industrial/Organizational Psychology, Overview ■ Job Stress ■ Stress

Further Reading

Arthur, A. (2000). Employee assistance programmes: The emperor's new clothes of stress management. *British Journal of Guidance & Counselling*, *28*, 550–559.

Ashforth, B. E., & Lee, R. T. (1997). Burnout as a process: Commentary on Cordes, Dougherty and Blum. *Journal of Organizational Behavior*, *18*, 703–708.

Beehr, T. A. (1995). *Psychological stress in the workplace*. London: Routledge.

Berridge, J., & Cooper, C. L. (1993). Stress and coping in U.S. organizations: The role of the employment assistance programs. *Work & Stress*, *7*, 89–102.

Cooper, C. L., & Marshall, J. (1976). Occupational sources of stress: A review of the literature relating to coronary heart disease and mental ill health. *Journal of Occupational Psychology*, *49*, 11–28.

Cooper, C. L., Dewe, P. J., & O'Driscoll, M. P. (2001). *Organizational stress: A review and critique of theory, research, and applications*. Thousand Oaks, CA: Sage.

Cordes, C. L., Dougherty, T. W., & Blum, B. E. (1997). Patterns of burnout among managers and professionals: A comparison of models. *Journal of Organizational Behavior*, *18*, 685–702.

Cox, T. (1978). *Stress*. London: The Macmillan Press.

Dewe, P., Cox, T., & Ferguson, A. (1993). Individual strategies for coping with stress at work: A review. *Work & Stress*, *7*, 5–16.

Dewe, P. J. (2001). Work stress, coping and well being: Implementing strategies to better understand the relationship. *Research in Occupational Stress and Well Being*, *1*, 63–96.

Ferguson, A., & Cox, T. (1997). The functional dimensions of coping scale: Theory, reliability and validity. *British Journal of Health Psychology*, *2*, 109–129.

Jex, S. M., & Beehr, T. A. (1991). Emerging theoretical and methodological issues in the study of work-related stress. *Research in Personnel and Human Resources Management*, *9*, 311–365.

Kahn, R. L., Wolfe, D. M., Quinn, R. P., Snoek, J. D., & Rosenthal, R. A. (1964). *Organizational stress: Studies in role conflict and ambiguity*. New York: John Wiley & Sons Inc.

Kahn, R. L., & Byosiere, P. (1991). Stress in organizations. In M. D. Dunnette, & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., pp. 571–649). Palo Alto, CA: Consulting Psychologists Press.

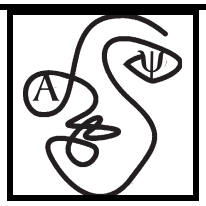
Lazarus, R. S. (1999). *Stress and emotion: A new synthesis*. London: Free Association Books.

McGrath, J. E. (Ed.) (1970). *Social and psychological factors in stress*. New York: Holt, Rinehart and Winston, Inc.

Maslach, C. (1982). *Burnout: The cost of caring*. Englewood Cliffs, NJ: Prentice-Hall.

Newman, J. E., & Beehr, T. A. (1979). Personal and organizational strategies for handling job stress: A review of research and opinion. *Personnel Psychology*, *32*, 1–43.

O'Brien, T. B., & DeLongis, A. (1996). The interactional context of problem-emotion and relationship-focused coping: The role of the big five personality factors. *Journal of Personality*, *64*, 775–813.



Jury Decision Making

Stephanie Albertson, Erin J. Farley, and Valerie P. Hans
University of Delaware, Newark, Delaware, USA

1. Introduction
2. Decision Making of Individual Jurors
3. The Psychology of Evidence
4. The Jury as a Group
5. Comparative Research on Juries
6. Conclusion
Further Reading

GLOSSARY

anonymous jury A jury for which the jurors' names are not revealed to the defendant or to the media. Used most often in high-profile trials or trials with security risks.

evidence-driven deliberation When jurors engage in general discussion of evidence in the case prior to voting, in contrast to verdict-driven deliberation, in which voting typically precedes evidence discussion.

hung jury A case in which the jury cannot agree on one or more verdict outcomes.

jury selection The process by which individual citizens are chosen for the jury.

scientific jury selection The use of scientific techniques by jury consultants to determine the individual demographic and attitudinal characteristics that are linked to views of a case. Techniques include community surveys, mock trials, focus groups, and in-court selection.

story model Theory of juror decision making in which jurors arrange evidence in the form of a narrative or story and match the story to the closest verdict alternative.

voir dire Questioning process during jury selection.

This article discusses trial by jury, including empirical jury research and applications. The use of the jury allows laypersons, rather than legally trained judges, to decide court cases. Critics question whether juries are able to decide cases competently and fairly. Addressing these criticisms, social scientists have examined the influence of individual factors and group variables on the jury decision-making process. Scholars have also studied the effects of jury reforms aimed at improving jury decision making. After summarizing the empirical findings and applied research, the article closes by discussing new directions in research on jury decision making.

1. INTRODUCTION

Trial by jury, in which citizens from the community hear evidence and decide on a verdict, is an important symbol of democracy. It allows peers, rather than legally trained judges or magistrates, to evaluate cases and reach binding decisions. As a vehicle for injecting lay participation into legal decision making, the jury has broad democratic appeal. Many contemporary legal systems incorporate juries, including those of the United States, Great Britain, Canada, Australia, New Zealand, and more than 40 other nations in Africa, Asia, South America, and the South Pacific. Some newly democratic countries, such as Spain, following the Franco regime, and Russia, after the breakup of the Soviet Union, have recently introduced or reintroduced the jury.

©2004 Elsevier Inc.
All rights reserved.

Public participation in the jury system is said to foster numerous important social and political purposes. First, it promotes accurate decision making that reflects community values. Close scrutiny of an individual case allows the jury to consider the specific facts as well as any extenuating circumstances, increasing the likelihood that the verdict will be fair and equitable. It permits citizens to engage in a key governmental function, educating them about the courts and law in the process. Juries also protect individuals against government oppression and provide a check on politically corrupt prosecutors and judges. Juries send signals through their verdicts about the community's standards and expectations. Thus, although juries today decide relatively few cases compared to earlier times, their verdicts remain important.

Although the jury system serves many positive functions, it has been subject to sharp criticism. The jury's competence has been attacked. Are juries able to comprehend the increasingly complex technical and scientific evidence in today's criminal and civil trials? Critics voice doubts concerning the jury's fairness, and express worry that verdicts are tainted with biases and prejudices of various kinds. For example, juries are viewed as overly sympathetic toward injured plaintiffs, yet at the same time as unduly harsh toward corporations and the wealthy. Jurors are described as "bleeding hearts" who acquit guilty defendants. Jurors are seen as too gullible in accepting illegitimate psychological defenses to criminal charges, such as the insanity defense, battered women's syndrome, battered child's syndrome, and post-traumatic stress disorder. In addition, there is also widespread sentiment in the United States that litigation has spiraled out of control, in part because of the promise of excessively generous jury awards to undeserving plaintiffs.

Debates surrounding the jury provide fertile ground for the application of psychological theory and research. Critically examining these debates and criticisms, psychologists have researched a wide range of jury topics, using mock jury simulation experiments, juror interviews, case analysis, and other methodologies. Psychological research on individual cognition, as well as the impact of juror demographic characteristics, attitudes, biases, prejudices, and personalities on verdict preferences and persuasion has uncovered important individual differences. This research has been applied to such areas as pretrial questioning, voir dire, and jury selection. Psychologists have also analyzed jurors' ability to integrate information, their perceptions and comprehension of evidence, and the dynamics of group decision making. Psychologists have used this research to propose specific jury reforms to trial procedure. In sum, the jury trial is a complex and fascinating institution that has been the site for advances in the field of applied psychology.

2. DECISION MAKING OF INDIVIDUAL JURORS

2.1. Individual Differences among Jurors

Psychological research has found a juror's decision to be the product of a complex set of factors. Not surprisingly, the trial evidence has the greatest impact. The evidence presented at a trial, and its strength and quality, constitute the prime determinants of jury verdicts. That finding has been replicated in both actual and mock jury studies time and again.

Nonetheless, research has also found certain other aspects of the trial that can affect juror judgments. In deciding criminal and civil cases, jurors draw on their life experiences such as crime victimization or physical injury. Jurors' own knowledge is brought to bear as they evaluate evidence in the case. Other factors that potentially impact individual cognition include a juror's social background, character, ideology, attitudes, and values. The nature of the evidence presented at trial, the way that attorneys structure their arguments, and the order in which witnesses testify can also influence evidence assessments.

2.1.1. Impact of Juror Demographic Characteristics

Although the strength of the evidence presented at a trial is the strongest predictor of the jury's verdict, many legal actors continue to believe that the juror's demographic characteristics are critically important. Oftentimes attorneys will exercise their peremptory challenges on the basis of characteristics such as race, ethnicity, occupation, gender, and religious preference. However, psychological research indicates that utilizing juror demographic characteristics is not a particularly useful or reliable method for predicting a prospective juror's propensity to favor one side over the other. The factors of age, gender, or race generally bear no relationship to verdict preferences, or have only slight predictive power. In fact, most individual differences among jurors typically account for only a small proportion of the variance in verdict outcomes. As a result,

eliminating or selecting prospective jurors based upon individual demographic characteristics may have little impact or might even backfire on attorneys.

Jury research on capital punishment is one of the few areas in which demographic variables are consistently related to punishment preferences. Women and racial and ethnic minorities are more likely than men and whites to oppose capital punishment. Even though strong opponents of the death penalty punishment do not sit as capital jurors, race and gender continue to play a role in capital jury decision making. Juries with a greater number of white men, for example, are more likely to recommend a death sentence than juries composed of more women and minorities. A juror's gender is also a reliable predictor in cases of rape and sexual harassment, with women generally giving more credence to the victim's testimony and evidence.

Demographic variables have low predictive ability in most cases. A prospective juror's knowledge, ideologies, beliefs, and experiences appear to be more significant than demographic characteristics in helping to explain why a juror reaches a particular decision.

2.1.2. Role of Juror Attitudes

Attitudes tend to be more reliable predictors of verdicts than demographic characteristics. Jurors' attitudes are influential in shaping their views and judgments of the parties, the attorneys, and the evidence presented at the trial. In turn, these assessments of evidence are related to the juror's verdict choices.

Psychological research on attitudes related to criminal jury trials has identified clusters of crime control and due process views. Those who endorse crime control beliefs and attitudes tend to have greater concern about crime, more distrust of defense attorneys, greater support for the police and prosecution, and more worry about defendants going free because of legal technicalities and illegitimate defenses. Crime control adherents tend to favor the prosecution in criminal trials. In contrast, those whose views reflect more due process concerns, such as worry about false convictions and support for the insanity defense and the exclusionary rule, are more likely to favor the defense. Research has also found a relationship between jurors' attitudes toward capital punishment and their verdicts about guilt or innocence. All else being equal, jurors opposing the death penalty are relatively more likely to favor acquittal than supporters of capital punishment.

In civil jury trials, found mostly in the United States, attitudes also play a role. Civil jurors frequently consider the plaintiff's claim with skepticism. They often question the severity of the plaintiff's claimed injury, sometimes blaming the plaintiff for the injury. Two psychological phenomena, the belief in a just world and defensive attribution, are linked to victim blame. In order to believe in a just and stable world, those observing an injury may minimize the injury and derogate the victim. To believe otherwise, that an innocent person is suffering, would challenge the prevailing belief that the world is predictable and fair. Psychologists have also theorized that when a victim is severely injured, observers are motivated to distinguish themselves from the victim and the threatening possibility of their own injury by engaging in defensive attribution. Defensive attribution may result in derogating the victim, blaming the victim for the injury, or minimizing the injury.

As for attitudes toward defendants in civil cases, civil jurors hold corporate defendants to a high standard of responsibility, and judge civil liability against this elevated standard. Respondents in several polls were found to have negative views of corporate management, and said they would be unable to be fair and impartial in cases involving tobacco companies and health maintenance organizations. However, whether general attitudes toward such businesses and organizations actually affect jury verdicts has not been established.

Pretrial publicity is another important factor that can help to shape initial views of trial evidence in a case. Exposure to pretrial publicity, particularly certain kinds of publicity such as criminal record information or emotional material, can influence the jury's decision making. Research has shown that negative pretrial publicity works in multiple ways: creating negative initial judgments of the defendant's character, shaping the processing of evidence, and influencing jury deliberations.

2.2. Psychology of Persuasion

The adversary trial is a two-way effort at persuasive communication. Psychologists as well as communication scholars have tested and applied theories of communication to the jury trial context. Expertise and trustworthiness are two dimensions of witnesses that influence how jurors view their credibility. Expertise is influential in creating attitude change, especially when the position advocated by the communicator is quite different from the recipient's initial attitude. Communicators arguing against their vested interests tend to be more persuasive compared to those arguing for their personal benefit; in addition, communicators with persuasive intent are viewed as less trustworthy.

An expert witness who is paid a sizeable amount of money to testify and does this for a living may be viewed by jurors as having low credibility compared to an expert witness who does not make a career from providing expert testimony.

Other factors found to increase persuasiveness include the communicator's attractiveness, similarity, and message comprehensibility. The more attractive the communicator, the more favorably he or she is viewed. Greater similarity to the observers (jurors) also leads to more positive appraisals. When it takes less effort to comprehend a message, individuals are more persuaded and are able to recall the information and arguments in the message to a greater extent. The degree of comprehension can become especially crucial during expert testimony, such as when scientific evidence is being presented. A juror's difficulty in comprehending complex evidence may reduce the juror's recall and hence the testimony's impact.

2.3. Applications

2.3.1. Voir Dire

During a typical jury selection, the judge and the attorneys may pose questions to prospective jurors during a period called the voir dire, to assess whether jurors can be fair and impartial in the upcoming case. The judge has unlimited ability to remove jurors for cause, if the judge determines that a prospective juror's views, attitudes, experiences, or connections to the parties or the case would make it difficult or impossible for the juror to be fair and impartial. In a number of countries, including the United States, attorneys may also remove a limited number of prospective jurors through peremptory challenges, without providing justification. Psychologists have studied the U.S. voir dire process and the effectiveness of attorneys' peremptory challenges. In past years, some attorneys systematically used their peremptory challenges in order to remove racial and ethnic minorities from the jury. The Supreme Court outlawed relying on race or gender in peremptory challenges, yet racial patterns in peremptory challenges persist.

From a psychological perspective, the voir dire questioning period is far from an ideal vehicle to assess the existence of bias and prejudice. Often the opportunity to question prospective jurors is severely limited. The questions may be close ended, in which only a yes or no response is called for, rather than open ended, allowing citizens to speak more fully about their views. This problem is compounded by the fact that it is usually clear which response is "right" or socially desirable. Even if the questioning is more expansive, it relies on jurors themselves to admit their own biases. It is difficult for jurors to become aware of their personal biases and admit them to a high status figure such as the judge.

Some scholars have drawn on psychological theory to propose and support recommendations for reforming the questioning of prospective jurors during the voir dire process of jury selection. More expansive voir dire should assist attorneys and the courts in their quest for impartiality. Incorporating both close-ended and open-ended questions during pre-trial questioning permits jurors to communicate their views and experiences in their own words, as opposed to answering with the less informative "yes" or "no." An individual rather than a group voir dire provides the judge and attorneys with more useful material for exercising challenges. These reforms would likely promote juror candor in disclosing pertinent information and discussing personal private experiences.

2.3.2. Jury Consulting and Jury Selection

Jury consulting is one tool used by lawyers to assist in identifying attributes of jurors that may negatively impact a case. Psychologists and other social scientists, operating as jury or trial consultants, have developed various techniques to assist in jury selection, collectively labeled scientific jury selection. Jury consulting is a fast-growing industry, with clients willing to pay large sums of money for this method of trial preparation assistance. It is largely an American phenomenon, employed only rarely in other countries.

Two prime methods used in scientific jury selection are telephone surveys and mock trials. In telephone surveys, members of the public are asked questions about their background characteristics, beliefs, attitudes, and views of the upcoming case. Analysis shows whether significant relationships exist between individual variables and case preferences. Jury consultants use these data to create profiles of favorable and unfavorable jurors.

Mock trials, ranging from the presentation of short statements of the evidence in front of focus groups to elaborate trial simulations, are also employed. Groups of community members observe the mock trial, discuss the evidence, reach a mock jury decision, and answer questions about the strengths and weaknesses of the case. The mock trial tests lay reactions to the evidence within the case and the arguments anticipated by the prosecution and defense. By analyzing mock jurors' reactions, the jury consultant can assist the attorney in preparing opening statements, trial presentation materials, and closing arguments.

Augmenting these methods are pretrial investigation of the jury pool, in-court jury selection assistance, witness preparation, and the use of "shadow juries" during the trial. A shadow jury consists of mock jurors who are chosen to mimic the actual jurors in the trial. They sit in court while the trial is going on, and periodically report their views of the evidence to the lawyers. In sum, research undertaken by jury consultants can assist attorneys in jury selection and the conduct of the trial, or even whether to go forward with a trial.

2.3.2.1. Limitations of Scientific Jury Selection The effectiveness of jury consulting is debatable. It has been argued that there is no convincing empirical evidence of the impact of scientific approaches to jury selection. Predicting jury decisions on the basis of telephone polls and mock juries is an inexact science. Critics point out the many differences between mock juries and real juries, who sit through the entire case, listen to all the evidence presented, hear the arguments by both sides, and have the responsibility for a final verdict. That is imperfectly captured by the empirical techniques that jury consultants employ. Many commentators believe that jury consulting about witness preparation and evidence presentation is likely to have more impact on trial outcomes than suggestions for jury selection.

2.3.2.2. Ethical Issues When scientific jury selection links jurors' demographic characteristics to trial evidence and arguments, ethical issues surface. First, any peremptory challenges based on the most obvious demographic characteristics, gender and race or ethnicity, contravene the U.S. Supreme Court's *Batson v. Kentucky* (1986) line of decisions. Categorizing individuals into particular groups based upon how they are predicted to behave or think establishes, enforces, and supports stereotypes and biases. A final ethical concern is related to the fact that scientific approaches to jury selection are expensive. In most cases, only wealthy clients can afford the full range of services. As a result, the disparity between the trial resources of the rich and poor is further aggravated.

3. THE PSYCHOLOGY OF EVIDENCE

3.1. Psychological Models of Juror Decision Making

The critical role of evidence in jury decision making has led to extensive study and debate. Many critics have charged that jurors misunderstand or incorrectly apply complex evidence in reaching their verdicts. To evaluate the merits of these criticisms, psychological research has examined the different ways that jurors comprehend and organize evidence.

Researchers have identified three distinct models of juror decision making: a Bayesian approach, an information integration model, and the story model. Each attempts to elucidate the process by which jurors interpret, evaluate, and combine evidence to arrive at a final verdict.

3.1.1. Bayesian Models

Bayesian models entail a formal approach to examining jurors' fact-finding process. This perspective focuses on how jurors update their beliefs about the likelihood of a certain event occurring based on the presentation of new evidence. As new evidence is presented throughout a criminal trial, jurors combine the pieces of evidence together on a single dimension that represents the probability of a defendant's guilt or innocence. Bayesian theory has been described as the "mental meter" model because it theorizes that jurors adjust their subjective probability of an event occurring as new evidence is presented. However, Bayesian models have been criticized for being too simplistic and for failing to capture the multidimensionality of thought processes during decision making.

3.1.2. Information Integration

The second popular model used to examine how jurors make decisions is the information integration model. Unlike Bayesian models, which are based on probability theory, the information integration model is an algebraic modeling approach. Information integration theory specifies how discrete information is combined into a final verdict. During trial, information may come from a number of different sources, including pre-existing biases of the jurors, jurors' observations of the defendant, the content of the testimony, the judge's instructions, and the attorneys' arguments. Jurors place a value and weight on each bit of important evidence. The evaluative information is integrated into a single judgment about the defendant's guilt or innocence.

3.1.3. Story Model

The story model posits that as jurors hear evidence, they actively construct it into a story or narrative. The

coherence of the story is shaped by the consistency, plausibility, and completeness of the evidence. The story may include character evaluations, motives, important events, and states of mind. Jurors actively fill in gaps in the testimony to complete the narrative through the incorporation of prior knowledge and preconceptions. When all these aspects of evidence are combined, they create a coherent story, and from this story jurors determine the appropriate verdict. Psychological research conducted on juror decision making has revealed a relationship between constructing a narrative and selecting a verdict that best fits the narrative account.

Importantly, jurors may ignore or reject information that is contradictory to their developing narrative or story. Jurors' rejection of evidence that is contrary to their expectations or individual beliefs is not only a critical component of decision-making models, but also a key to understanding and improving how jurors perceive evidence. For example, evidence may be presented in a way that is difficult to understand or inconsistent with expectations, resulting in jurors disregarding it and focusing on more understandable evidence. However, if legal actors are aware of this potential, the presentation of evidence can be altered to increase juror understanding and comprehension.

3.2. Perceptions of Evidence

The presentation and comprehensibility of different types of evidence also merit investigation. Jury perceptions of three types of evidence have been examined: general evidence, eyewitness evidence, and expert evidence.

3.2.1. General Evidence

In terms of general evidence, one of the consistent concerns among critics is jury comprehension. Critics charge that the jury is unable to understand simple and complex trial evidence. However, research studies of jury decision making are largely at odds with this conclusion. The classic early study of jury decision making, conducted by Kalven and Zeisel (1966) in the 1950s, examined judicial opinions about jury verdicts in thousands of trials. The study revealed that the majority of judges agreed with the jury's verdict in the case. When judges disagreed with jury verdicts, they did not do so more frequently in difficult cases, suggesting that case complexity was not the source of the disagreement. A high rate of judicial agreement with jury verdicts has also been found in subsequent studies.

Nonetheless, concerns about jury competence have persisted. Research on jury performance confirms that even in fairly complex cases, jurors are generally able to understand enough of the evidence to reach a competent verdict. Research on comprehension of legal instructions, however, is more negative about the jury's ability.

The timing of evidence or legal instructions has been examined for its impact on juror decisions. Traditionally, the judge gives legal instructions to the jury at the end of the trial, right before deliberations. This was based on the assumption of a recency effect, in which immediate past events are remembered better than more distant events. On the other hand, the primacy effect posits that early information has great impact. Jurors usually form some opinions about the trial before the closing remarks by attorneys. Therefore, the primacy effect suggests that giving jurors instructions at the beginning of the trial would have greater impact. Introductory legal instructions at the start could help jurors assess legally relevant evidence throughout the rest of the trial. In addition, applied research shows that jurors who receive instructions at the beginning of the trial report greater satisfaction with the process and a better ability to evaluate evidence according to the proper legal principles.

3.2.2. Eyewitness Evidence

Eyewitness identification is another example of evidence that may have a significant impact on jury decision making. Although eyewitness identification is perceived as the most influential evidence in jury decision making, research has revealed it to also be one of the most unreliable forms of evidence. Studies have shown that an eyewitness's confidence is only weakly associated with an accurate identification; however, the greater an eyewitness's confidence, the more likely the jury will perceive the identification to be accurate. In addition, the status of the eyewitness has an effect on jurors' beliefs about witness credibility. For example, jurors are more likely to give weight to police officers' identifications than those of regular citizens, believing that training promotes police officers' accurate identifications. However, research reveals that police officers are no better at eyewitness identification than regular citizens.

Psychological research reveals important flaws in eyewitness testimony and the types of procedures that police use to confirm eyewitness identification. Yet, many of these flaws are not common knowledge among the jurors. Because of the importance of eyewitness testimony, the courts have come to recognize

the value of informing jurors about them so that eyewitness identification is allocated an appropriate weight. Psychologists have helped to reform police identification procedures, and now routinely testify as experts in jury trials, educating juries about the research findings concerning eyewitness evidence.

3.2.3. Expert Evidence

The use of expert testimony is another important factor affecting juror decision making. The use of complex evidence during trials has greatly increased within the last few decades. As a result, there has been an increasing reliance on expert witnesses to explain the evidence. Therefore, the question of how juries handle expert evidence on scientific and technical topics has become of great interest.

It has been claimed that expert witness credentials overly impress jurors, resulting in too much weight being given to their testimony. On the other hand, juror comprehension of expert testimony is also criticized on the theory that if it is too complex jurors will simply dismiss it. Also, critics and legal actors fear that permitting expert testimony will result in opening the door to trials becoming a battle between expert witnesses.

Research has revealed that jurors do have difficulty comprehending complex evidence. However, lack of juror comprehension can often be attributed to inadequate presentation of evidence and testimony, not the inherent inabilities of jurors. Improving instructions and evidence presentation will in turn increase juror comprehension. In addition, research has cast doubt on a number of assumptions that depict jurors as rubber stamps who passively accept expert testimony. According to empirical studies, instead of being unduly influenced by experts, jurors critically and adequately evaluate both the credentials of the experts and the quality of their testimony. Finally, even if some complex evidence is beyond the ability of the average juror, case studies indicate that jurors can usually understand enough of it to be able to arrive at a justifiable verdict.

3.3. Applications

3.3.1. Promotion of Juror Competence through Trial Reform

Understanding how jurors comprehend and utilize different types of evidence has led to a number of proposed reforms of the traditional adversary trial, several of which have been adopted by multiple jurisdictions. Adhering to the traditional idea of jurors as passive observers has contributed to jurors' difficulty in comprehending evidence and trial procedures. One general reform is first to accept jurors as more active participants in the trial process, acknowledging the fact that jurors are not blank slates and that they begin to assimilate and integrate evidence long before the end of the trial.

To maximize juror comprehension and decision making, reformers suggest permitting jurors to take notes during trial, allowing jurors to ask questions of witnesses through the judge, simplifying the language used in legal instructions, giving instructions at the beginning versus the end of the trial, giving written instructions to each juror, and giving jurors multipurpose notebooks to better organize large amounts of evidence and information during complex trials.

Many judges and attorneys are apprehensive about trial reform and concerned about changing traditional aspects of the adversary jury trial. Yet, some of these fears, when tested, have proved to be unfounded. For example, the idea that juror note taking distracts jurors, causes them to miss important information, and offers an advantage to jurors who are better note takers is not borne out. Research shows that jurors are not distracted and do not miss important evidence due to note taking. In fact, note taking encourages jurors to stay alert during trial, serves as a very good memory aid, and increases juror satisfaction with the trial process.

4. THE JURY AS A GROUP

4.1. Group Decision Making: Theory and Research on Group Performance

4.1.1. Decision-Making Process

During deliberation, the door to the jury room is locked. No one other than the jurors is present. Therefore, what transpires behind the closed door of the jury room has been, until fairly recently, something of a mystery. Researchers have acquired information on the deliberation process by interviewing actual jurors following their jury duty and also through mock jury experiments.

Once a foreperson or jury leader has been selected, typically one of two types of discussion follows. In an evidence-driven discussion, jurors first take up a general discussion of the evidence in the case. Jurors attempt to develop a joint account of the events related to the trial. Once the evidence has been thoroughly discussed, the

group takes a vote about the verdict preferences of individual jurors. In contrast, in a verdict-driven discussion, jurors begin by taking a vote at the start of deliberations. Discussion is subsequently organized around jurors' votes (e.g., innocent or guilty) and evidence is typically discussed in terms of supporting one side or the other. Another dynamic that occurs during deliberation is the tendency for jurors to align themselves with others holding similar positions.

Psychological research has found the amount of group participation to be associated with individual characteristics. For example, men typically speak more than women during deliberation. Individuals with more education and higher-status occupations tend to dominate group discussion. The foreperson is also quite active in the deliberation process and may serve as an influential voice during deliberation. Studies have shown that most juries have several nonparticipating jurors during the deliberation process. They speak rarely during the group discussion and have little impact on the group; their only contribution may consist of taking part in voting. As a result, during deliberations the jurors who dominate discussions have more potential to influence the views of other jurors. At the same time, jurors who withdraw from critical discussion may be less influential in deliberations.

In most deliberations, the first vote, or ballot, will often reveal what the verdict will be. The majority opinion is usually the final verdict. However, minority votes can sometimes prevail. In criminal trials, they are often more successful in persuading the majority to overturn their votes when arguing for an acquittal as opposed to conviction. Minority jurors can also influence the majority in civil trials. If the jury is asked to award damages, minority jurors can be influential in creating a compromise on the verdict or award amount.

Jury deliberation serves critical functions. Research shows that it promotes jury competence by providing the opportunity for a collective pooling of information and the correction of mistaken memories or conclusions. In addition, deliberation also clarifies and solidifies jurors' positions. The consensus about the evidence and the verdict that typically emerges also increase the legitimacy of the jury's eventual decision in the case.

4.1.2. Persuasion during Deliberation

Social psychologists have used the vehicle of jury deliberation to examine social influence and persuasion processes in small groups. In turn, theoretical advances have been applied to jury phenomena. Theories about group phenomena, including conformity effects and group polarization, have been examined in the context of the jury deliberation process. Conformity effects of the majority's initial opinions have been documented in jury deliberations, where the initial vote distribution is a strong predictor of the jury's final verdict.

Informational influence and group polarization found in mock jury deliberations suggest that juries operate in much the same fashion as many other task groups. Social psychological theories about group process have been used to assess the impact of different structural variables related to the jury such as composition, size, and decision rule, as discussed below.

4.1.3. Hung Juries

Many factors can impact why a jury deadlocks, or is "hung," including individual, group, evidentiary, and institutional factors. Individual factors could be incompetence, attitudes, and lack of social skills. An ineffective foreperson or group leader can impede deliberation, for example, by not permitting everyone to speak, allowing one or two jurors to dominate the deliberation, and ignoring arguments by minority jurors. An ineffective decision-making process, by balloting at the start of deliberations as opposed to thoroughly discussing the evidence prior to voting, can also influence vote outcome. Evidentiary factors are also significant; the ambiguity of the evidence can influence the propensity for a jury to deadlock.

Jury research has found that members of hung juries tend to vote earlier than members of verdict juries. Their deliberation appears to be verdict driven, as opposed to evidence driven. When juries do hang, group dynamics are at work. There may be a conflict on the jury, there may be unreasonable jurors, or one or two jurors may dominate the deliberation. Jury members may not be open minded, and may spend too much time and effort attempting to persuade and convince fellow jurors. Numerous factors act and interact with one another to influence jury verdict outcomes, from individual characteristics to group dynamics.

4.2. Applications

Understanding the jury decision-making process in terms of group interaction has important implications and applications for real-world situations. For example, psychological research has examined the impact of decreasing jury size, altering the decision rule

requirement for juries, allowing jury anonymity, and improving jury decision making through jury reforms.

4.2.1. Impact of Jury Size and Decision Rule

The jury's size and its decision rule have gone through a number of controversial changes. With respect to jury size, what was once long assumed to be a constitutional guarantee, a trial by a 12-member jury, was dismantled by a number of important U.S. Supreme Court cases during the early 1970s. The U.S. Supreme Court decided the number 12 was a "historical accident," finding that state court juries of six did not violate the Constitution. Decreasing the requirement from a 12- to a six-member jury was seen as a way to decrease expenses, save time, and increase efficiency at no cost to the jury trial system.

However, psychological research has identified a number of negative consequences of reducing jury size. The larger the jury size, the more likely it is to be representative, and in turn the more likely it is to reflect community views. In reverse, the smaller the jury size, the less likely it is that minorities in the community will be represented on the jury. In addition, the unpredictability of jury verdicts increases as jury size decreases. The smaller juries used in civil trials are more likely to result in a wider range of jury awards.

Furthermore, six-person juries appear to deliberate differently than larger juries. Larger juries have more people to rely upon to recall evidence with greater accuracy, which decreases the likelihood of relying on incorrect information. Jurors in larger groups are more likely to be diverse in terms of backgrounds, experiences, and opinions. This diversity of perspectives increases the overall quality of jury deliberations and increases the validity of the final verdict. One juror within a six-member jury proposing an opposing position is in a significantly weaker position compared to two jurors in a twelve-member jury.

In addition to changing jury size requirements, countries have modified the long-held requirement that juries must be unanimous in their verdict. During the 1970s. the U.S. Supreme Court decided that unanimous juries in state courts were not constitutionally mandated. The United Kingdom switched from requiring unanimity to permitting majority verdicts as well. These changes were again based on unsupported assumptions about how juries deliberate. Subsequent research shows that eliminating jury unanimity has negative effects on the quality of juror deliberations. For example, jurors with majority rule verdicts instead of unanimous verdicts are more likely to be verdict driven than evidence driven. Although verdict-driven deliberations may occur in jury trials that require a unanimous verdict, they are more likely to occur in majority rule juries and are more likely to decrease the quality of the deliberation process.

Jurors deliberating under a unanimity rule requirement are more likely to be evidence driven, and evidence-driven juries are more likely to have longer and more thorough deliberations. A unanimity requirement increases the likelihood that all viewpoints will be appropriately addressed during the decision-making process. Majority rule juries are more likely to be verdict driven, resulting in deliberations that are less thorough and significantly shorter. Verdict-driven deliberations are shorter and less thorough because once the majority requirement is reached, there is no need to listen to the remaining minority opinion, resulting in a termination of deliberations. The quality of deliberation decreases because minority opinions do not get as much attention; when they are addressed, they are not sufficiently discussed. As a result, it should not be surprising that jurors on non-unanimous juries report less satisfaction with the process and the verdict, while jurors on unanimous juries report greater satisfaction with both the process and the final verdict.

4.2.2. Anonymous Juries

Another controversial technique used to improve the decision-making process is juror anonymity. Juror anonymity means withholding the identities of the jurors from the public, the media, and the attorney. There are a number of reasons why judges have felt it necessary to order that the jury members be anonymous: to prevent corruption, to decrease fear of criticism from unpopular decisions, and to decrease fear of retaliation or harassment in high profile cases.

Many critics have claimed that jury anonymity decreases juror accountability, the ability of lawyers to adequately investigate juror bias during voir dire, and public confidence in the fairness of jury trials. However, some of these concerns can be managed. With anonymous jurors, lawyers have access to the same questionnaires used in regular trials and also maintain the ability to question jurors; the only difference is that jurors are identified by numbers instead of names. Preserving the questioning process during voir dire, which uncovers important background and attitudinal information of a potential juror, seems more important for creating an unbiased jury than knowing the juror's name.

Juror anonymity is also viewed as improving the overall deliberation process. For example, jurors in high-profile

cases may fear retaliation or harassment from the defendant, the lawyers, the media, or the public. As a result, maintaining anonymity may increase general participation from the population and also increase the quality of deliberations, due to decreased fear of retaliation. Anonymous jurors may be more comfortable discussing unpopular or controversial topics, resulting in an increased richness of deliberations. Also, anonymous juries may be freed from the pressure of following popular opinion in high-profile trials. There are many potential benefits to anonymous juries, psychological and otherwise, that could increase the quality of the overall decision-making process, and in turn increase the fairness of jury verdicts. They deserve systematic examination and study.

4.2.3. Active Jury Reforms

Psychological research indicates that modifying jury size and decision rule may produce some negative consequences for jury decision making. In addition, psychological research has suggested that other jury reforms may improve the overall trial process and jurors' individual and group decision making. One controversial reform proposition is based on the idea of active jurors, allowing jurors to discuss evidence throughout the trial, instead of waiting until the end of the trial as is commonly the case. Permitting trial discussions is based on the belief that jurors form opinions about evidence and testimony throughout the trial, and discussing evidence and questions as they arise will help dispel inaccurate interpretations or beliefs and improve the quality of final deliberations.

The state of Arizona is one place where the concept of the active juror has been vigorously pursued. Following substantial jury reforms, Arizona now permits juror note taking, question asking, and trial discussions, all with the aim of improving juror competence. Research studies of some of these reforms show strong support among jurors and judges, but less enthusiasm for the reforms among attorneys. Attorney fears about distraction and prejudgment, however, have not been borne out in systematic studies of the operation of the reforms.

5. COMPARATIVE RESEARCH ON JURIES

Much of the empirical jury research to date has taken place within the United States, England, and Canada, studying how the jury within a single country operates. Yet, countries that have adopted the jury have molded the institution to fit their unique circumstances and cultures. Researchers have only recently begun to take a comparative approach, analyzing how different jury systems function.

Examining how different procedures in various countries affect jurors' decision-making ability is an emerging area for psychological research. For example, although both the United States and England use adversarial models of procedural justice, the degree of involvement of judges and lawyers varies significantly. English judges, who are permitted to comment upon the evidence, may be more influential in the typically more civil and cooperative British jury trial, while U.S. juries may be more distracted by the greater adversariness of the typical American trial.

Another important question that may be answered through comparative research is how juries measure up to other methods for injecting lay sentiment into law, such as lay judges, lay magistrates, and mixed tribunals, which combine lay judges with law-trained judges. One key finding is that when lay judges decide cases together with law-trained judges, the legally trained judge tends to dominate, minimizing the input of lay perspectives. In contrast, the jury decides in isolation, relatively protected from the direct influence of the law-trained judge. This new comparative work, now in its infancy, holds promise for understanding the consequences of different types of lay participation systems.

6. CONCLUSION

The jury represents an important vehicle for the expression of public perspectives on justice. Critics of the jury system would prefer to weaken or silence the public's voice. Yet psychological research offers empirical evidence that has revealed the basic soundness of jury decision making while simultaneously pointing toward effective reforms of the jury system. Acknowledging jurors as active yet imperfect actors in the trial process allows legal practitioners and researchers to propose jury trial reforms that have the ability to improve the presentation of evidence and maximize the quality of deliberations to not only increase the satisfaction of jurors, but also reinforce the legitimacy of their final verdict.

See Also the Following Articles

Decision Making ■ Eyewitness Testimony

Further Reading

Abramson, J. (2001). *We, the jury: The jury system and the ideal of democracy*. Cambridge, MA: Harvard University Press.

Dann, B. M. (1993). 'Learning lessons' and 'speaking rights': Creating educated and democratic juries. *Indiana Law Journal*, *68*, 1229–1279.

Diamond, S. S. (1990). Scientific jury selection: What social scientists know and do not know. *Judicature*, *73*, 178–183.

Hans, V. P. (2000). *Business on trial: The civil jury and corporate responsibility*. New Haven, CT: Yale University Press.

Hans, V. P., & Vidmar, N. (1986). *Judging the jury*. Cambridge, MA: Perseus.

Hastie, R. (Ed.). (1998). *Inside the juror: The psychology of juror decision making*. Cambridge, UK: Cambridge University Press.

Horowitz, I., & Bordens, K. (2002). The effects of jury size, evidence complexity, and note taking on jury process and performance in a civil trial. *Journal of Applied Psychology*, *87*, 121–130.

Kalven, H., Jr., & Zeisel, H. (1966). *The American jury*. Boston, MA: Little, Brown, & Co.

Munsterman, G. T., Hannaford, P. L., & Whitehead, G. M. (Eds.). (1997). *Jury trial innovations*. Williamsburg, VA: National Center for State Courts.

Vidmar, N. (Ed.) (2000). *World jury systems*. Oxford: Oxford University Press.

Labor Turnover

Peter W. Hom
Arizona State University, Tempe, Arizona, USA

1. Introduction
2. Identifying Who Quits
3. Theories about Why Employees Quit
4. Methods for Controlling Turnover
5. Conclusion
Further Reading

GLOSSARY

ecological fallacy The assumption that the relationship between two variables is the same across different levels of analysis (e.g., individual vs organizational).
employee turnover Voluntary cessation of organizational membership by members who receive monetary payment for their participation and labor.
expectation-lowering procedure Orientation for new recruits that deflates initial job expectations without providing a realistic preview about a particular job.
fit A component of job embeddedness that represents employee compatibility with the company or environment.
image violation An experienced event that violates personal ideals or aspirations, inducing thoughts of quitting.
job avoidance Less extreme forms of avoiding the workplace.
job embeddedness A composite construct representing the web of connections that keep employees from leaving.
job satisfaction The favorability of an employee's attitude toward the job.
links The formal and informal relations that an employee forms with members in the organization and community.
matching script A preexisting plan to exit the job once some preconditions are met.
met expectations The alignment between the prehire job expectations of beginning employees and their later work experience that affects their socialization to the job and retention.
organizational commitment Employee identification with the values and goals of the employing institution and feelings of loyalty toward that institution.
realistic job previews Full communication of the positive and negative aspects of the job to new entrants during hiring or orientation to the job.
sacrifices A component of job embeddedness that refers to material or psychological benefits that an incumbent forfeits when he or she vacates the job.
shocks Jarring events that trigger thoughts of quitting.
weighted application blank A detailed application form that is converted into a selection test to screen out job applicants with a high propensity to quit.
work–family interference Organizational interference with employees' ability to participate in family or extraorganizational activities.

Employee turnover is the voluntary termination of membership in employing institutions that is initiated by employees. Over the years, scholars have identified which employees tend to sever their employment ties, why they do so, the process by which they leave, and how they might be prevented from quitting. Although important strides toward the understanding and prediction of turnover have been made, several areas of inquiry remain unresolved or unexplored.

©2004 Elsevier Inc.
All rights reserved.

1. INTRODUCTION

Employee turnover is the voluntary cessation of membership in an organization by an individual who receives monetary compensation for participation in that organization. This phenomenon has long been an inquiry of investigation for industrial and organizational psychologists, with more than 1000 studies having appeared over the years. This article briefly reviews the theoretical and practical significance of this behavior, investigations of the identities of leavers, theory and research about the causes of employee turnover, and the prevention and control of this behavior.

Turnover attracts the attention of organizational scientists and employers alike because this highly visible behavior may signal poor quality of working life and ineffective organizational processes. In particular, turnover often imposes significant financial costs to firms. When employees quit, employers must expend funds to find and train replacements. To illustrate, human resources professionals project that the cost of one turnover incident ranges between 93 and 200% of the leaver's salary. Apart from direct financial costs, turnover may engender broader costs. For example, excessive personnel attrition may impair customer service and/or harm patients. Moreover, many chief executive officers of leading corporations, such as Microsoft and General Electric, believe that the flight of top talent jeopardizes firm performance. Finally, U.S. corporations worry that excessive attrition among minorities and women may hamper their efforts to diversify their workforces, especially top management teams.

2. IDENTIFYING WHO QUITS

2.1. Demographic and Personality Correlates of Turnover

Turnover researchers have long sought to identify who quits. According to a comprehensive meta-analysis, Griffeth and colleagues observed that employees who are prone to quit generally are young and single, are not married to working spouses, have no (or few) children, and/or have short tenure in the organization. Contrary to popular opinion, this meta-analysis detected no racial or gender differences in quit rates. Nonetheless, these conclusions should be regarded with caution. Griffeth and colleagues' meta-analysis included only seven studies on the race–turnover relationship and combined all racial minorities together (ignoring racial differences in quit rates) while failing to examine subgroup differences among women (e.g., educated women and uneducated women quit at different rates). In comparison, the Attrition and Retention Consortium, which is composed of 25 Fortune 500 corporations that share quit statistics on managers and professionals for benchmarking, found that women quit more than do men. Based on a combined sample of 81,346 racial minorities, this consortium further observed that Hispanic and African Americans leave at higher rates than do European and Asian Americans during the early period of employment. Because such findings may expose consortium participants to litigation or embarrassment, such data are not available from the academic literature (and so their omission biases demographic meta-analyses).

In addition to demographic traits, turnover researchers have examined other personal attributes of leavers. Specifically, a meta-analysis of the "Big Five" personality dimensions reported that long-tenured employees tend to be conscientious, agreeable, and open to experience. Supporting the old maxim that the "best predictor of future behavior is past behavior," it was further demonstrated that employees with a history of unstable employment exhibit a greater risk of quitting.

2.2. Performance–Turnover Relationship

By far, the personal characteristic attracting most turnover studies is job performance because whether high or low performers tend to quit greatly determines the costliness of turnover. Organizations may even welcome personnel attrition if those who are exiting are their less valuable or less valued job incumbents. Testing the performance–turnover relationship, several meta-analyses have concluded that substandard performers are more prone to quit, whereas outstanding performers are more stable.

Despite such impressive consistency, some researchers have witnessed inverted U-shaped turnover–performance relationships, such that superior performers (because they have greater employment opportunities) and poor performers (because they are unhappy or face impending dismissals) terminate employment more often than do moderate performers. Conventional wisdom had long assumed that greater flight among top performers occurs only in occupations where achievements are publicly known (e.g., professional athletes, university scholars). Because their professional accomplishments are visible

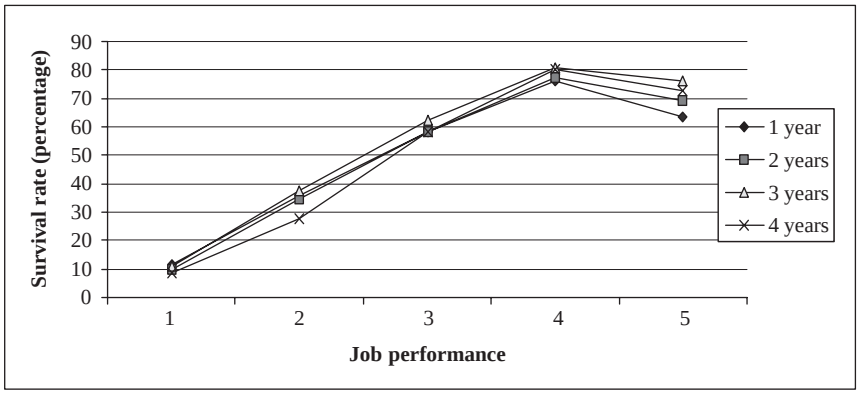


FIGURE 1 Job survival rates as a function of job performance and firm tenure.

signals to the external marketplace of their ability, superior performers can readily secure alternative employment elsewhere (including receiving more unsolicited job offers from outside employers). Recently, a few studies have detected that this trend toward greater job mobility among productive incumbents is more common, reporting higher turnover rates among high-performing managers and Swiss bankers (Fig. 1). As a result, these recent findings qualify traditional observations that low performers exit voluntarily at greater rates than do high performers. When professional accomplishments are externally visible or objectively verifiable (e.g., fast-track promotion history for managers), curvilinear (rather than inverse) relationships between performance and turnover may well emerge.

3. THEORIES ABOUT WHY EMPLOYEES QUIT

3.1. Theories about Intermediate Linkages between Job Dissatisfaction and Quits

More than 45 years ago, March and Simon pioneered a theory about why employees quit that continues to dominate current perspectives. They conceptualized two prime turnover determinants—"ease of movement" (e.g., job opportunities that allow employees to quit) and "desirability of movement" (e.g., job dissatisfaction that motivates employees to leave)—that persist in recent theorizing. Extending this line of thinking, Mobley first elaborated the various stages of the withdrawal process. Specifically, he theorized that job dissatisfaction prompts employees to begin thinking about quitting. Dissatisfied incumbents next evaluate their prospects of obtaining employment elsewhere and the costs of quitting (e.g., loss of seniority benefits). If attractive alternatives exist and turnover costs are low, disgruntled employees form intentions to search for other jobs and actively pursue other work. After securing other job offers, these employees compare them with their current positions. If this comparison favors an alternative, they will ultimately quit.

Since its inception, Mobley's model had dominated theory-testing research throughout the late 1970s and 1980s, yet empirical tests failed to sustain Mobley's causal structure. Instead, in 1991, Hom and Griffeth validated an alternative sequence of structural linkages among withdrawal antecedents (Fig. 2).

In 2001, Hom and Kinicki validated an extended version of this model (Fig. 3) with a national retail store sample. This version additionally considered local unemployment conditions, work–family conflict (i.e., job interference with employees' ability to attend to family and/or extraorganizational pursuits), and job avoidance (i.e., less extreme forms of avoidance such as absenteeism and tardiness).

In this theory, interrole conflict increases dissatisfaction and withdrawal cognitions, whereas unemployment conditions directly lower quit rates. Furthermore, job dissatisfaction increases job avoidance, which in turn augments withdrawal cognitions.

3.2. Structural Determinants of Turnover

Apart from theoretical accounts of how dissatisfaction translates into exits, Price and Mueller introduced another framework that also initiated a long research

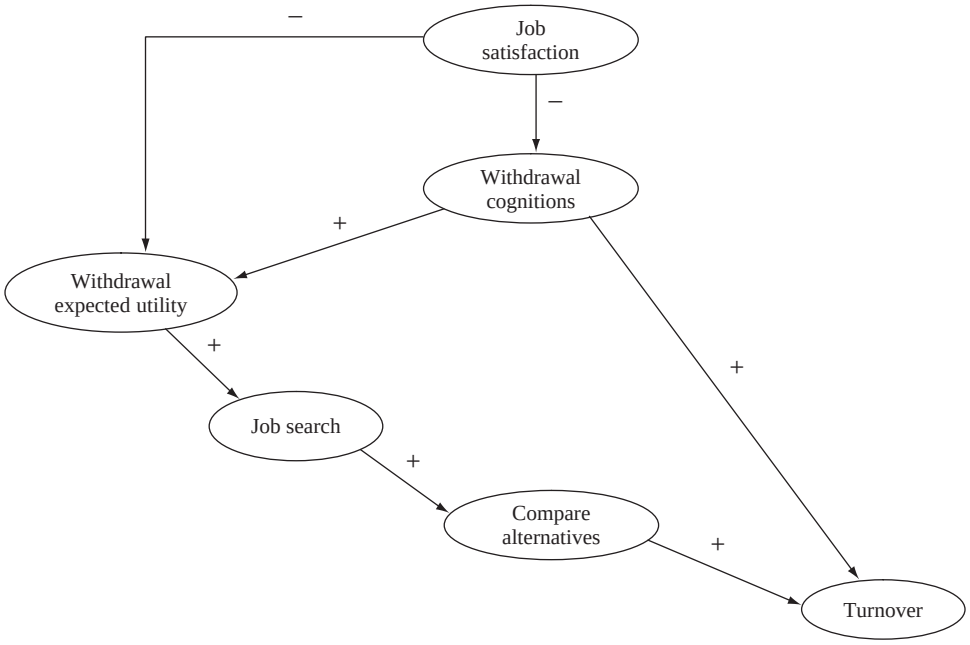


FIGURE 2 Hom–Griffeth model of dissatisfaction-initiated quits.

stream. Unlike psychologists (who focus on micromediation among psychological constructs), these sociologists identified a broad array of more distal structural antecedents of turnover (e.g., characteristics of the work setting, patterns of social interaction within the workplace) as well as influences external to the workplace (e.g., occupational commitment, family obligations). Price and Mueller theorized that role overload, routine work, centralized decision making, work-related communications, social integration (i.e., friendships at work), pay, fair reward allocation, and promotional opportunities affect job satisfaction, whereas

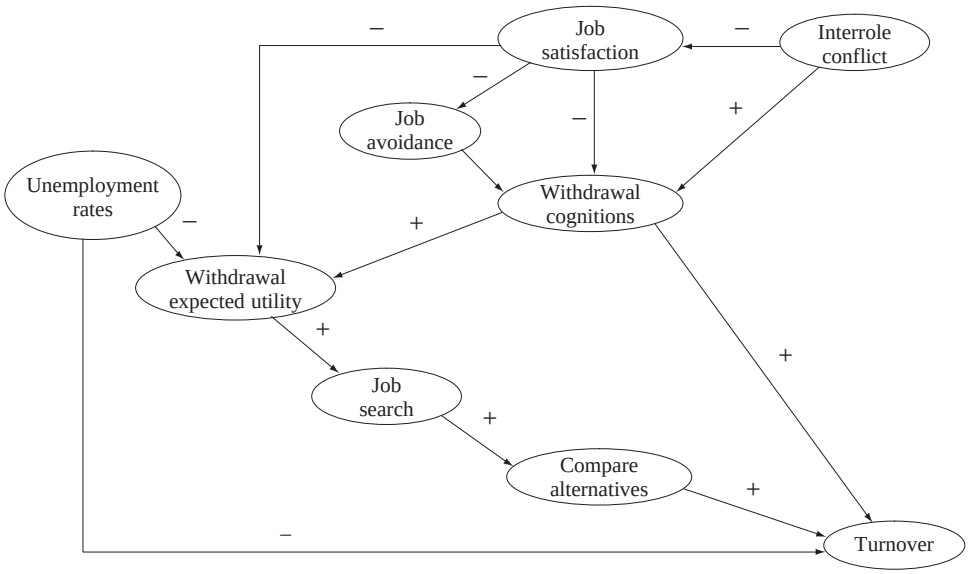


FIGURE 3 Hom–Kinicki model: Expanded theory of dissatisfied quits.

professionalism (i.e., commitment to one's professional standards and norms), general training, and "kinship responsibility" (i.e., family responsibilities and relatives living nearby), as well as job satisfaction, underpin organizational commitment (i.e., firm identification and internalization of firm values). Commitment and perceived job opportunities, in turn, stimulate the decision to quit, and this decision, as well as perceptions of job opportunities, affects turnover. Although studies have not validated every pathway in the Price–Mueller model, studies have broadly established that structural determinants affect either job satisfaction or organizational commitment—or both—and that environmental antecedents affect quits (although whether their effects are direct or indirect via commitment or quit intentions remains indeterminate).

3.3. Unfolding Model

Marking a radical departure from traditional perspectives, Lee and Mitchell proposed a provocative new theory to overcome the shortcomings of these viewpoints. They argued that conventional thinking wrongly assumes that job dissatisfaction is the prime impetus for quitting and that prospective leavers quit only for other employment. Prevailing schools of thought also presume that prospective leavers evaluate alternatives using a rational analysis of their costs and benefits. Instead, Lee and Mitchell's "unfolding model" posits that nonattitudinal causes—namely "shocks" or jarring events that stimulate thoughts of quitting—instigate quits more often than do attitudes toward the job (Fig. 4). Their formulation also holds that some leavers exit without jobs in hand (including, e.g., exiting the workforce to return to school) and that leavers often make quick "compatibility" evaluations of alternative jobs rather than subject them to rational subjective expected utility analyses.

The unfolding model further proposes that leavers follow one of several pathways when exiting. Some leavers follow Decision Path 1, which is precipitated by a shock (typically some expected personal event). This shock causes incumbents to probe their memory for a "matching script" or preexisting action plan, and if one is detected, this can lead to automatic withdrawal. For example, a woman may quit when she becomes pregnant (the shock) because she had already planned that she would stay at home to raise her first child on becoming pregnant. Other leavers follow Decision Path 2 because they experience a shock that triggers an "image violation." Lee and Mitchell defined three types of images: value image (e.g., personal

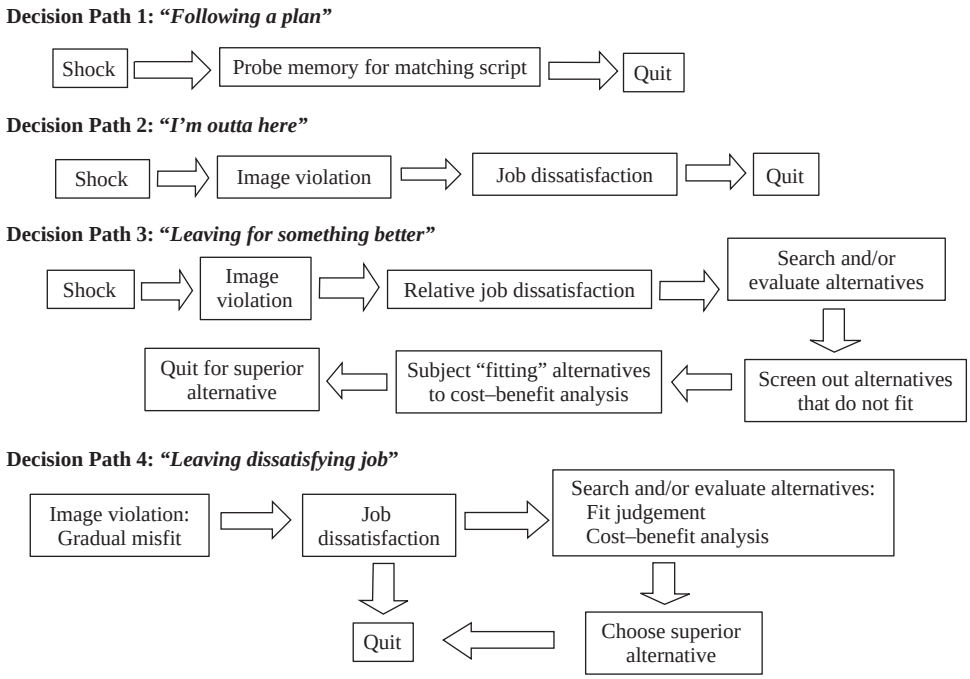


FIGURE 4 Lee and Mitchell's unfolding model.

values), trajectory image (e.g., personal goals), and strategic image (e.g., behavioral strategy toward goal attainment). In this turnover path, a shock violates employees' images and causes them to feel dissatisfied and to directly quit. For example, a hospital may shift to team-based nursing (the shock), which violates a nurse's image because the nurse values individual-based nursing. The nurse becomes unhappy and leaves the hospital without engaging in a prior job search.

Still other leavers follow Decision Path 3, which is initiated by some unsolicited job offers or inquiries (the shock). This shock leads job incumbents to compare these new jobs with their current ones and leads to "relative" dissatisfaction if the alternatives are superior (i.e., incumbents experience "opportunity costs" by staying). The prospects of "greener pastures" elsewhere motivate incumbents to pursue and evaluate other alternatives (including the unsolicited jobs). They first screen alternatives using a quick compatibility test to determine fit with their personal images, rejecting those that lack such a fit. Next, the job seekers engage in more deliberate cost–benefit analyses among alternatives meeting the fit test and ultimately quit to take the "best" alternative. To illustrate this path, Lee and colleagues described a nurse who received a job offer from a private clinic (the shock). She had been frustrated for a while but had not contemplated leaving until someone offered her a better position. She then investigated this clinic position thoroughly before exiting the hospital.

Finally, some leavers take Decision Path 4 (regarded by Lee and Mitchell as the conventional path, similar to that portrayed by the Hom–Kinicki model). Here, leavers gradually become disenchanted with their jobs as they progressively experience image violation. Some may quit as soon as their dissatisfaction reaches some threshold. This is what Jim Manzi, chief executive officer of Lotus Development, did when IBM bought his company. Manzi became unhappy with the direction and vision for the newly merged companies and left without a new job in hand. Alternatively, some leavers who take Decision Path 4 seek and judge other employment options and exit once they secure better positions.

Using qualitative methodology, empirical tests of the unfolding model interview former employees to delineate the routes by which they withdrew from their previous jobs. More than 92% of leavers followed one of the four decision paths depicted by the unfolding model. Moreover, the studies employing these tests report that 25% of leavers quit without searching for employment elsewhere and that shocks prompt more turnover than does job dissatisfaction. Finally, 39% left due to external shocks such as pregnancies, unsolicited job offers, and spousal relocations. Despite these provocative findings, retrospective reports, which may be distorted by postdecisional dissonance and faulty recall, provide inadequate validation for the unfolding model. Further verification of this theory requires predictive longitudinal designs—a standard methodology for verifying turnover models.

3.4. Job Embeddedness

In 2001, Mitchell and colleagues proposed explanations for why people stay. Existing formulations explain why people leave but fail to account for why people stay, possibly implicating different psychological processes. Mitchell and colleagues introduced "job embeddedness" as a composite construct representing the web of connections that keep employees on the job. Job embeddedness is composed of three dimensions: links, fit, and sacrifice. Links represent the formal and informal connections between employees and institutions and others. Thus, friendships at work and in the community, as well as extended family in the community, may keep employees from abandoning their workplaces. Fit represents employee compatibility with the company and environment. If employees have values compatible with those of the workplace, enjoy recreational and entertainment activities in the community, and belong to local community organizations (e.g., churches), they may stay. Finally, sacrifices correspond to the costs of material and/or psychological benefits that are forfeited by resignations. Thus, employees stay if they have nonportable job benefits (e.g., pensions), seniority-based benefits, and community advantages (e.g., live in safe neighborhoods close to work, have prime seating at sporting events). Testing this construct, Mitchell and colleagues documented that job embeddedness accounts for unique variance in the prediction of turnover that is not explained by conventional predictors.

4. METHODS FOR CONTROLLING TURNOVER

In contrast to the substantial explorations of turnover theories, examinations of ways in which to reduce turnover remain limited. In 2001, Griffeth and Hom discussed the more promising approaches that have been supported empirically.

4.1. Employee Selection

The most widely validated selection tool is the weighted application blank (WAB), that is, a detailed application form converted into a selection test. To develop a WAB, an employer would compare application blank responses of former employees (leavers) and current employees (stayers). When these groups answer a particular question differently (e.g., stayers might report higher incomes on their last jobs than do leavers), a scoring key is generated for that question and "scores" the leavers' response (low wage) higher than the stayers' response (high wage). Scoring keys are developed for all questions that discriminate between stayers and leavers. The sum of the special scores for all items yields an overall score for quit propensity. When used for hiring, the WAB essentially compares job applicants' answers on the application blank with those of leavers (former employees). If certain applicants closely resemble leavers in their application responses, these applicants would receive high WAB total scores and would be disqualified from hiring. Despite the simplicity and proven success of WABs, few employers use them because questions about certain background attributes may disproportionately reject minority or women applicants (and prompt discrimination lawsuits) or may constitute an invasion of privacy.

In addition to assessing background attributes to identify applicants at risk for quitting, organizations might consider personality testing. However, conventional personality scales are susceptible to faking by job seekers who (mis)portray themselves in a favorable light to gain employment. To overcome such social desirability, Bernardin suggested a personality-like measure of quit propensity using a forced-choice format. In this approach, respondents choose a descriptor (about themselves) from a pair of descriptors that are matched on equal discomfort, although only one is a valid answer. To create this test, employers must interview current employees and their superiors to identify discomforting aspects about a job (e.g., "your schedule for work changes from day to day"). They then write statements based on these events (questionnaire items). Next, employers also write statements about discomforting situations that are irrelevant to the job (e.g., "buying a new television and later learning that it goes on sale next week"). Then, a set of judges is recruited to rate the discomfort levels of relevant and irrelevant statements. Then, employers pair one relevant statement with an irrelevant statement that is matched for discomfort. Once many sets of paired statements are identified, they are included in the final personality test. When taking this test, prospective employees would select a statement from each pair depicting an event most distressing to them. If they choose the "valid" statement (representing an actual negative job experience) from a given pair, they would earn a "higher" score for that answer. The test would indicate that applicants are a poor fit for a job if they accumulate high test scores (based on summing their responses to all statement pairs). Put differently, these individuals would likely be upset by discomforting events that are part of the job. They would ultimately become dissatisfied and quit. Bernardin and colleagues established that this personality inventory predicts turnover among customer service representatives and motion picture theater personnel.

Finally, selection interviews have potential for identifying prospective leavers if they are properly designed and executed. Notably, Schmidt and Rader described a promising structured telephone interview. To develop this screening device, employers must carry out in-depth interviews with outstanding performers in a particular job to identify their behavioral tendencies and traits (i.e., "themes"). Next, potential interview questions assessing each theme are written out as potential interviewees would be asked how they exhibit such traits. During the next step, two groups of employees—30 high performers and 30 low performers—are interviewed with these questions individually, and their remarks are transcribed. Questions differentiating between high and low performers are retained, and a scoring guide that credits interview responses indicative of the presence of high-performer themes is created. During telephone interviews, job applicants would be asked to respond to these questions (and to provide examples). Applicant responses would also be recorded and later scored by trained analysts to derive scores that assess how well applicants exhibit the high-performer themes. Summarizing multiple validity studies, Schmidt and Rader estimated that this interview predicts employment stability moderately accurately (r = .39).

4.2. Interventions for Reducing Turnover

Turnover scholars have most validated realistic job previews (RJPs), that is, comprehensive profiles of a job—portraying its positive and negative aspects—that are communicated to new hires during hiring or orientation. Meta-analyses conclude that RJPs modestly reduce exits among beginning employees in a wide range of

jobs (e.g., nursing, military) with a variety of media (e.g., booklet, audiovisual). To account for the effectiveness of job previews, theorists have often cited met expectations and self-selection as underlying mechanisms. According to met expectation theory, RJPs reduce reality shock by decreasing the gap between newcomers' initial (often inflated) expectations about a new job and subsequent work experience. RJPs clarify what the job entails and increase the odds that newcomers' expectations will be met on the job, thereby improve retention. The self-selection mechanism posits that RJPs help job applicants to make more informed job choices, in turn helping them to better decide whether or not their needs would be satisfied by the job. Those who self-select themselves for the job (despite the RJPs) are better suited for the job and will likely stay, whereas those who are "misfits" will decide not to accept the job (and avoid becoming future turnover casualties). Despite the popularity of these explanatory schemes, empirical corroboration of RJP explanations remains limited.

In 1998, Buckley and colleagues introduced the expectation-lowering procedure (ELP), which reduces job expectations without providing actual information about the job. After all, expensive job analyses are needed to properly design RJPs, and they may lose their efficacy if job content changes. To offset these limitations, Buckley and colleagues designed a special orientation for new hires, who are instructed in the theory of met expectations (and are explained the causes and consequences of unmet expectations) and share their own personal experiences with reality shock. In an experimental test, Buckley and colleagues showed that the ELP reduced turnover among new assembly line workers as much as did an RJP.

Finally, turnover scholars have validated job enrichment—increasing the intrinsic rewards of a job by expanding skill variety and empowering incumbents—as another means to promote job loyalty. Indeed, a meta-analysis concluded that job enrichment is more effective than are RJPs for promoting retention, supporting Wanous's maxim that RJPs cannot substitute for good jobs.

5. CONCLUSION

Organizational studies have greatly expanded our understanding of the turnover phenomenon. Nonetheless, much theoretical and empirical work remains. Specifically, empirical verification of turnover formulations has lagged theoretical advances. Conventional static designs typically assess turnover causes at one time and use them to predict turnover at a later time. This standard methodology overlooks the implicit causal assumption inherent in most theoretical frameworks that changes in turnover's antecedents foreshadow its occurrence. To better corroborate causality, more rigorous evaluations of prevailing theories should apply latent growth modeling or hierarchical linear modeling to longitudinal data (i.e., multiple waves of predictor assessments).

Moreover, conceptual models of the withdrawal process have been validated primarily in North American samples. Yet American transnational firms increasingly employ global workforces and face the daunting challenge of how to bind their workers to employment. Because retention strategies based on ethnocentric domestic theories may falter abroad, it is imperative that researchers validate turnover models (and loyalty-inducing practices) in various cultural settings. Furthermore, theorists have narrowly focused on explaining why individuals quit and derive organization-wide practices for reducing aggregate quit rates from individual-level findings. Yet inferring company-wide prescriptions from individual-level data and theories risks committing the "ecological fallacy." Accordingly, future research should develop (and test) formulations that account for why rates of personnel attrition vary across organizations.

See Also the Following Articles

Conflict within Organizations ■ Organizational Culture and Climate ■ Person–Environment Fit ■ Personnel Selection ■ Traits

Further Reading

Bernardin, H. J. (1987). Development and validation of a forced choice scale to measure job related discomfort among customer service representatives. *Academy of Management Journal*, *30*, 162–173.

Buckley, M. R., Fedor, D. B., Veres, J., Wiese, D., & Carraher, S. M. (1998). Investigating newcomer expectations and job-related outcomes. *Journal of Applied Psychology*, *83*, 452–461.

Griffeth, R., & Hom, P. (2001). *Retaining valued employees*. Thousand Oaks, CA: Sage.

Griffeth, R., & Hom, P. (2004). *Innovative theory and empirical research on turnover*. Greenwich, CT: Information Age.

Griffeth, R., Hom, P., & Gaertner, S. (2000). A meta-analysis of antecedents and correlates of employee turnover: Update, moderator tests, and research implications for the next millennium. *Journal of Management*, *26*, 463–488.

Hom, P. W., & Griffeth, R. (1991). Structural equations modeling test of a turnover theory: Cross-sectional and longitudinal analyses. *Journal of Applied Psychology*, *76*, 350–366.

Hom, P., & Griffeth, R. (1995). *Employee turnover*. Cincinnati, OH: South-Western.

Hom, P., & Kinicki, A. (2001). Toward a greater understanding of how dissatisfaction drives employee turnover. *Academy of Management Journal*, *44*, 975–987.

Lee, T., & Mitchell, T. (1994). An alternative approach: The unfolding model of voluntary employee turnover. *Academy of Management Review*, *19*, 51–89.

Lee, T. W., Mitchell, T. R., Wise, L., & Fireman, S. (1996). An unfolding model of voluntary employee turnover. *Academy of Management Journal*, *39*, 5–36.

March, J., & Simon, H. (1958). *Organizations*. New York: John Wiley.

Mitchell, T., & Lee, T. (2001). The unfolding model of voluntary turnover and job embeddedness: Foundations for a comprehensive theory of attachment. In B. Staw, & R. Sutton (Eds.), *Research in organizational behavior* (Vol. 23, pp. 189–246). Oxford, UK: Elsevier Science.

Mitchell, T., Holtom, B., Lee, T., Sablynski, C., & Erez, M. (2001). Why people stay: Using job embeddedness to predict voluntary turnover. *Academy of Management Journal*, *44*, 1102–1121.

Mobley, W. (1982). *Employee turnover*. Reading, MA: Addison–Wesley.

Price, J., & Mueller, C. (1986). *Absenteeism and turnover of hospital employees*. Greenwich, CT: JAI.

Schmidt, F., & Rader, M. (1999). Exploring the boundary conditions for interview validity: Meta-analytic validity findings for a new interview type. *Personnel Psychology*, *52*, 445–464.

Villanova, P., Bernardin, J., Johnson, D., & Dahmus, S. (1994). The validity of a measure of job compatibility in the prediction of job performance and turnover of motion picture theater personnel. *Personnel Psychology*, *47*, 73–90.

Wanous, J. P. (1980). *Organizational entry: Recruitment, selection, and socialization of newcomers*. Reading, MA: Addison–Wesley.

Wanous, J. P. (1992). *Organizational entry* (2nd ed.). Reading, MA: Addison–Wesley.

Leadership and Culture

Roya Ayman
Illinois Institute of Technology, Chicago, Illinois, USA

1. Introduction
2. Culture
3. Leadership Research across Cultures
4. New Topics in Culture and Leadership
5. Summary and Conclusion
Further Reading

GLOSSARY

accommodating strategy When one party in a conflict gives all of his or her rights for another's to win.
collaboration strategy When two parties focus on the problem and the best solution in which both can win.
competitive strategy When one party in a conflict is focused on him- or herself to win at all cost.
compromising strategy When parties negotiate a settlement.
consideration Behavioral style that is friendly, open to input, respectful to others, etc.
contingency approach Model of leadership that describes leadership in context.
derived etic approach The coalescence of theories or measures about a concept that are developed in two different cultures.
emic approach A method of studying culture theories, models, and measures as evolved from inside a culture.
endogenous researchers Researchers who conduct research within their own culture.
etic approach A method of studying culture in which existing theories, models, and measures in one country or culture is validated in another.
exogenous researchers Researchers who conduct research outside their own culture.
global leaders Leaders who perceive complexity and are inclusive, respecting people of all cultures equally, and who are concerned about what is the best solution in the long run.
idealized influence Behavioral styles that provide vision and unify the team behind a common goal.
imposed etic approach Approach using existing theories or measures from one culture in another culture with almost no modification other than translations.
individual consideration Style of behaviors such as attentiveness to diversity of needs and values of team members and addressing each on their own basis.
initiating structure Behaviors that focus on getting the job done, such as goal setting, clearing expectation, criticizing wrong behavior, rewarding good behavior, planning, etc.
inspirational motivation Behavioral style that exudes energy and emotion in the actions to be taken.
intellectual stimulation The process of getting team members to think critically.
international leaders Leaders who have the tendency to impose their own cultural values and policies on others.
leader–member relationship The leader's certainty that the team is united and supportive of him or her.
position power The leader having a title, and the legitimacy and resources to reward and punish team members.
situational control A leader's amount of control and influence in a given situation.
task structure A leader's knowledge of how to get information or conduct a task.
withdrawal When parties to a conflict leave the relationship.

©2004 Elsevier Inc.
All rights reserved.

Globalization of work economy has made the interplay of culture and leadership an important topic of consideration. This article covers the topic of leadership and culture, demonstrating a symbiotic relationship between the two. It is shown that the definition and methods used in the study of culture in leadership can affect the results. Culture as determined by economico-political conditions had the most impact in determining differences in leadership across cultures. Accordingly, managers in countries that are similar in terms of industrialization and modernization have more similarity in their values and even their behaviors. Through the analysis of studies on behavior of leaders across cultures, it became evident that the definition of leadership is highly cultural dependent. More specifically, the meaning of the behavior, the effect of the behavior, and the level of the behavior varied across cultures. One outstanding finding was that Asian cultures do not seem to consider task-related and people-related behaviors as two separate leader behaviors. Additionally, more recent issues of leadership, such as working in multicultural settings and global leadership, are addressed.

1. INTRODUCTION

Globalization is an inevitable and challenging evolution in today's society. Leaders in any social and economic institution are at the helm of this evolution of multicultural interdependence. Therefore, a deep understanding of the concept of leadership within cultures, across cultures, and globally is of paramount importance. To gain an appreciation of the interaction between the leader and culture, consider the following anecdote: An Asian manager in the United Kingdom. noticed that one of his employees was absent from work due to ill health. The manager visited the employee at home, bringing flowers and a box of sweets, as was customary in the manager's culture. In his mind, he was showing individualized consideration and social support for his employee. Nevertheless, the employee, acting according to his own cultural norms, filed a complaint, charging the manager with distrust and invasion of his privacy. The issue here is not who was correct, but the elucidation of the role that culture plays in interpreting behaviors.

The preceding anecdote illustrates the fact that leadership is embedded in culture, and that the two have a symbiotic relationship. To explore this, this article presents a brief definition of culture and the cross-cultural methods. Next, different leadership definitions and the role of culture in their universality is discussed. The topics of leadership as seen in some countries, the leader in cultural exchange, and global leadership conclude this presentation of culture and leadership.

2. CULTURE

2.1. Definition of Culture

"Without men, no culture, certainly; but equally, and more significantly, without culture, no men." Clifford Geertz made this powerful if not precise definition of culture, highlighting the importance of culture in the study of human beings. Scholars concur that culture refers to a constellation of a group's collective values, expectations, attitudes, and behaviors. In addition, the symbiotic relationship between culture and humans is interactive. While cultural environments model humans, humans also affect the culture. In other words, culture refers to an all-encompassing and long-lasting depiction of a human social entity.

Regardless of the definition of culture used, in cross-cultural psychology, researchers focus on two distinctions of culture: cultural values and physical characteristics. Cultural values primarily include (1) collectivism and individualism, (2) power distance, (3) uncertainty avoidance, (4) masculinity or industriousness, (5) Confucianism, and (6) time orientation. Physical parameters include country of origin or residence, as well as physical attributes of individuals, such as gender and color.

2.2. Cross-Cultural Methodology

Because results of cross-cultural research are dependent on the method used to study leadership, this section defines the various alternatives to and factors in cross-cultural research. The methods used to incorporate culture into psychological research have evolved since almost 50 years ago. Traditionally, two types of approaches have been recognized: (1) when existing theories, models, and measures in one country or culture are validated in another, or the etic approach, and (2) when theories, models, and measures evolve from inside a culture, or the emic approach. The etic approach is further expanded to include imposed etic and derived etic. Imposed etic refers to using existing theories or measures from one culture in another with almost no modification other than translations. Derived etic, on the other hand, refers to an

approach in which a theory or measure developed in two different cultures coalesces; it is that which is common between the two cultures.

In addition to the different approaches to cross-cultural research, the researcher's cultural background can impact the formation of the research question and the interpretation of results. Therefore, cross-cultural researchers can be categorized by their level of involvement with a culture as either exogenous or endogenous investigators. Exogenous researchers are those who conduct research outside their own culture. Therefore, they are not entangled with the culture, and high degree of objectivity is the outcome. Danger, however, lies in the fact that exogenous researchers' assumptions may be derived from their own culture or their stereotypic view of the culture they are studying. Endogenous researchers, those who conduct research within their own culture, may be less objective but are more sensitive to nuances and meanings of their own culture.

Thus, cross-cultural research usually starts with emic measures in one culture, which then can be tested in the other cultures as imposed etic measures. The items that best describe both cultures will form the derived etic measure. Most research on leadership thus far can be categorized as imposed etic, particularly the validation of theories developed in the United States in other cultures. Additionally, U.S. scholars have conducted most of the research and developed most of the main theories of leadership; hence, the existing knowledge about leadership is heavily influenced by North American values and expectations. Therefore, given the extent to which exogenous researchers from one specific culture have dominated the theories and conceptualizations of leadership, it is mostly representative of one perspective. For this reason, multicultural research teams that include endogenous researchers are the new trend and the recommended form of cross-cultural research.

3. LEADERSHIP RESEARCH ACROSS CULTURES

Concern for leadership has a long history. In China in 5000 BC, Confucius advised the Chinese emperors; in 12th-century Persia, Unsur-u'l-Ma'ani wrote an essay advising the Ziari rulers, and in 15th-century Europe, Machiavelli wrote *The Prince*, guiding the rulers of the time. In the late 19th century, Carlyle formulated the Great Man theory, later known as the leadership trait approach, which had the most impact on the majority of research on leadership in the United States.

Initially, leadership was approached with a focus on either the leader alone, or the leader as part of a social context. When the attention was focused on the leader, his or her traits and values were considered as well as his or her leadership behaviors. The models of leadership that describe leadership in context are referred to as contingency approaches. In the following sections, the etic approaches to leadership highlight the effect of culture on the universality of the leader's traits and behaviors and the validity of contingency approaches to leadership across cultures.

3.1. Etic Approach: The Leader's Characteristics

The trait approach to leadership is the oldest approach that initially incorporated both dispositional characteristics, including values, attitudes, and traits (e.g., intelligence, locus of control, confidence, optimism, and adaptability), and physical characteristics (e.g., height and attractiveness). More than two-thirds of the 20th-century research on leadership focused on identifying dispositions that best describe a leader. Intelligence, adaptability, adjustment, aggressiveness, dominance, independence, originality, self-confidence, and optimism were dispositions that most studies identified as separating leaders from non-leaders.

Cross-cultural comparison has most often focused on managers' attitudes and values. Two approaches are used to examine the effect of culture on managers' similarity of values. One approach examines the values of managers in different countries; the other compares managers from a similar country of origin living in different countries.

When different nationalities were compared, it accounted for approximately 28% of the variance in managers' attitudes. In contrast, when the effect of similarity among managers from Australia, India, Japan, Korea, and the United States was assessed, and nationality versus industrialization was compared, it was the closeness in the level of the societies' industrialization, language, and traditionalism that accounted for most of the leaders' value similarity, not the country they lived in. Although managers from Australia and the United States shared similar values, the Japanese and Korean managers did not show a similarity of values. Such differences were attributed to varying levels of industrialization of the two countries. Therefore, it is not the country or nationality per se, but the proximity in

socio-economic conditions of those countries that determines similarity between managers/leaders.

To isolate the impact of culture from national origin, researchers have contrasted the effect of social/ethnic culture and country of residence. In these studies, the values of managers from various countries were compared to the values of managers who lived in the United States. The result showed that those managers who came from countries that were at similar levels of modernization as the United States. had similar values. So, for example, the U.S. managers of both Caucasian and Chinese background had values more similar to those of managers from Hong Kong and more different from those of Mexican and Mexican American managers. This result not only illustrated issues related to difference in industrialization between the United States and Mexico, but it also alluded to the potential reason that Mexico's adaptation to industrialization is relatively more difficult.

Likewise, the results of a study that compared managers with Chinese heritage from three different countries (the People's Republic of China [PRC], Singapore, and the United States) showed that the values of the Chinese managers who lived in Singapore and the United States were quite similar to each other, and were quite different from those who lived in the PRC. Again, the role of modernization and industrialization of the country of residence seems to contribute to the similarity between the leaders' values.

In short, the aforementioned studies illustrated that industrialization and education are powerful societal forces in unifying managers' values across countries. No one specific set of values defined all managers. However, the conditions in which the managers functioned did have an impact in making them more similar.

The majority of research and theories on leader behavior in the United States relies heavily on subordinate perception of the leader's behavior. The content of these behaviors has been as specific as conflict management strategies and influence, and as broad as transactional and transformational leadership. All the various theories have reported some cross-cultural validity in predicting outcomes and some level of universality in the presence of these behaviors.

To review this aspect of leadership and the role of culture, first the behaviors are briefly defined. Subsequently, three cross-cultural concerns that have been presented in testing the universality of these behaviors are discussed. These concerns are (1) comparative level of occurrence of these styles across cultures, (2) difference in behaviors determining the styles across cultures, (3) their functionality across cultures.

Theorists and researchers have identified many leader behaviors with a plethora of empirical research.The oldest leadership behavioral styles studied in U.S. research are consideration and initiating structure. Consideration refers to behaviors that are friendly, open to input, respectful to others, etc. Initiating structure includes behaviors that focus on getting the job done, such as goal setting, clearing expectation, criticizing wrong behavior, rewarding good behavior, planning, etc. These were identified as two distinct categories of leadership behavior that help a group meet its goals of maintaining harmony and accomplishing tasks successfully.

Subsequently, leadership researchers acknowledged that these are transactional-type behaviors, and they recommended that effective leaders go beyond the status quo. Thus, the transactional and transformational leadership paradigm was presented. Behaviors included in transactional leadership style were active or passive monitoring of performance and giving feedback, also the act of giving rewards when due. Transformational leadership included four behavioral styles, intellectual stimulation, individual consideration, inspirational motivation, and idealized influence. Intellectual stimulation is getting team members to think critically. Individual consideration includes styles of behaviors such as attentiveness to diversity of needs and values of team members and addressing each on their own basis. Inspirational motivation consists of behavioral styles that exude energy and emotion in the actions to be taken. Finally, idealized influence includes behavioral styles that provide vision and unify the team behind a common goal. Theorists such as Bass and Avolio have been explicit in clarifying the relationship between the transformational and transactional styles. The issue is not that the leader should engage in one or the other style, but rather that transformational style augments the effects of transactional style. Overall, men and women leaders do not differ greatly in these behaviors; however, women leaders are perceived to be slightly more participative than men leaders. Additionally, women leaders are perceived to be slightly more transformational and to give more contingent rewards as compared to men leaders.

Conflict management and influence tactics are significant leadership behaviors; many scholars define leadership as a process of influencing others to do what they would not do otherwise. Rahim's conflict management model identifies two adversaries and the strategies they may take in managing the conflict between them. When one party is focused on winning at all costs, this is known as competitive strategy. When one party gives all his or her rights for another's to win, this is

considered accommodating strategy. If both parties leave the relationship, it is known as withdrawal. If the parties negotiate a settlement, this is a compromising strategy. Finally, if the two parties focus on the problem and the best solution in which both can win, this is referred to as collaboration strategy. Influence tactic are behaviors identified empirically when people were asked how they influence others. The tactics most commonly supported in U.S. research were rational persuasion, ingratiation and coalition tactics, inspirational and personal appeals, exchange, and assertiveness.

Overall, each of these behavioral paradigms in the United States has received support across settings and populations. That is, both the reliability of the measures assessing the styles and the functionality of the styles in relation to organizational outcomes have received empirical support. Next, the effect of culture on these behaviors is examined.

The first question that many cross-cultural researchers have had when using these theories across cultures is, Do all managers behave similarly across cultures? Empirical studies have found that the magnitude of the manifestations of these styles vary across cultures. This was established by comparing the mean scores of the frequency of occurrence of the various styles in different countries. For example, it was found that New Zealander managers were described as more transformational than transactional, whereas the Taiwanese managers were seen as behaving equally in both transformational and transactional styles. It was also reported that in a manufacturing telecommunication transmission equipment industry, the U.S. managers, in comparison with the German managers, were described as more charismatic and were reported as being more inspirational when motivating their employees.

Some cross-cultural leadership researchers became aware that specific behaviors that define a style might vary from culture to culture. For example, it is reported that in Indonesia for leaders to be able to inspire and motivate their subordinates, they must first persuade their subordinates that they are competent leaders, so leaders boast about themselves. This behavior, on the other hand, is not acceptable in Japan or in some other countries. In another study, comparing Taiwanese responses to those of Americans, it was found that the value system of the respondent had an effect on which specific behavior was related more to transformational or transactional leadership. Also, in a recent multicultural study, it was found that the behaviors contributing to each dimension of transactional and transformational leadership styles varied when the respondents were not culturally homogenous. However, in comparing U.S. managers to those from Bangladesh, the results showed that the five-conflict management styles, as defined in the United States, were also valid for Bangladesh. Therefore, the meanings of the conflict management styles is the same across these cultures.

In regard to universal functionality of leader behavior, we can examine two studies assessing the relationship between leader behavior and organizational outcomes. A five-country (Japan, South Korea, Taiwan, Mexico, and the United States) comparison illustrated that charismatic, supportive, and contingent rewards had a positive impact across cultures on employee work-related attitudes, whereas participative, directive, and contingent punishment in relation to these organizational outcomes varied by country.

In an experimental study in the United States, two confederates acted as either a transformational leader or a transactional leader, and they both led a group of Asian Americans or Caucasians. It was found that the Asian groups were more collectivist while the Caucasian groups were more individualistic; the collectivist groups (Asian) generated more ideas when the leader was a transformational leader than when the leader was transactional. However, the individualistic groups (Caucasian) generated more ideas with a transactional than with a transformational leader. These studies demonstrate that the team members' cultural values and experiences moderate the functional relationship between transactional and transformational leadership and organizational outcomes. Examining the effect of culture on the effect of influence tactics cross-culturally, two cross-cultural studies with managers from Hong Kong, the PRC, Switzerland, and the United States were examined. The findings supported the fact that cultural values affected the effectiveness of managers' choice of influence tactics. The Western managers, as compared to the Chinese managers, rated direct task-oriented influence more effective, and they rated influence tactics that were based on personal relations less effective. More specifically, Schmidt and Yeh quoted the Chinese saying, "Give benefit, but show authority." This quote is a good example of the impact of a society with high power distance and high collectivist cultural values.

The cross-cultural evidence shows that although the notions of the U.S.-driven concepts of leadership are present in all cultures, they do not manifest at the same level or in the same way, or have similar effects on employees. Therefore, it is necessary to recognize that

the cultural context in which a leader is functioning will impact the behavioral choices he or she makes; it also impacts the interpretation of their behavior. Ultimately, culture moderates the effectiveness of these leader behavior styles. On the other hand, the mangers from cultures with similar level of socio-economic development seem to be more similar compared to those from societies with various level of development. Thus, it is possible that as the gap between level of education and advancement shrinks, the similarity among leaders across cultures will also increase.

3.2. Etic Approach: Contingency Approach to Leadership

Three main contingency models of leadership are recognized in leadership research: Fielder's contingency model and leadership effectiveness, Mitchell and House's path-goal theory, and Vroom's leadership decision-making theory, which has the most empirical research. Because the path-goal theory of leadership has received mixed support, this article focuses on the role of culture in two models: Fiedler's contingency model of leadership effectiveness and Vroom's normative decision-making theory.

3.2.1. Contingency Model of Leadership Effectiveness

The contingency model of leadership effectiveness is a leadership model that combines the leader's trait and its match with the situation in predicting leadership effectiveness. The results of past research have found support for the model across various industries and populations. The situation in this model is defined by assessing the leader's amount of control and influence in the situation, referred to as situational control. In high and low situational control conditions, leaders who are more task-motivated than interpersonal are more effective; whereas in moderate situational control, leaders who are more relationship-motivated rather than task-motivated are more effective.

Fielder identified three factors that contribute to situational control. In order of importance, they are the leader–member relationship, the task structure, and the position power. The leader–member relationship refers to the leader's certainty that the team is united and supportive of him or her. Task structure refers to the leader's knowledge of how to get information or conduct a task. Position power refers to the leader having a title and the legitimacy and resources to reward and punish team members.

The trait of the leader is assessed with the least preferred coworker (LPC) scale. This construct assesses the leader's motivational orientation, that is, whether the leader gives priority to task accomplishment or to people's feelings. High score is indicative of relationship orientation and low score is indicative of task orientation.

The model has been tested in across multiple cultural settings, and in many it has shown to be successful in predicting a leader's effectiveness. This model is sensitive to cultural milieu, as the culture of the situation can be assessed by the situational control factor in the model. Thus, the model has not only gained cross-cultural support, it has also demonstrated that it can incorporate culture as part of its assessment of the situation.

The LPC score, depicting the leader's trait in this model, has a normal distribution in the United States with a mean of 65. The range of leaders' LPC scores varies across cultures. For example, the average score for Mexican managers on this scale was higher than the U.S. samples, and the average for Iranian managers was higher than both the United States and Mexico. On the other hand, the managers from England, Thailand, and the United States were not different on their scores on LPC scale. Also, no gender difference was found when comparing men's and women's LPC scores. Furthermore, the distributions of the score vary across cultures. For example, almost all respondents from Navaho nation (a Native American tribe) scored so high on this scale that it was not possible to identify any individual with low scores (task oriented).

The model, therefore, allows for the assessment of the interference of culture on leadership effectiveness, and it has shown varying degrees of cross-cultural validity. However, due to varying ranges and distribution of the LPC scores, it could be concluded that across the respondents from the United States, Mexico, Iran, and Navaho nation, the U.S. managers tend to have the prepotency to give higher priority to task accomplishment than to their relationships compared to the managers from other cultures. Nevertheless, the premise that task-motivated leaders do better when they feel in control, and relationship leaders do better when engaging team members, seems to be valid across cultures.

3.2.2. Normative Model of Decision Making

Based on group dynamic and decision-making principles, this prescriptive model of leadership decision-making

strategy was developed. The model states that leaders typically are confronted with five decision-making strategy options that range from very authoritative, to consultative, to participative. This model assists the leader in assessing the situation and choosing the most effective decision-making strategy. The situation in this model is assessed through answers to questions on the leader's knowledge, the team's receptivity to the decision made, and the time available to make the decision. In brief, when leaders have the information and are not dependent on the subordinates to carry out the decision, they can choose more autocratic decision-making styles and spend less time on this process. However, if the leader either does not have all the facts or is dependent on the subordinates' support, he or she needs to take more time and use more consultative and participative styles of decision making.

The validity of the model was tested on both men and women leaders. It was found that male leaders who were perceived as autocratic were evaluated as moderately positive. In contrast, female leaders who were perceived as autocratic were rated negatively.

Cross-cultural studies using this model showed that the choice and success of decision-making style is dependent on politico-economic conditions. For example, when this model was tested in Austria and Poland, it was found that Polish managers had a harder time following the model's prescription to use participative decision-making styles when necessary. This became more evident as the decision became more important. Furthermore, results showed that only after the fall of communism in Poland did Polish managers show a trend toward more participative practices. Therefore, these two studies demonstrated the impact of politico-economic condition on the leader's decision-making choice. Also, when the ideal leader's decision-making style was compared across the three German-speaking countries of Austria, Germany, and Switzerland, the results showed that the managers in all three countries preferred a team-oriented leader. However, the Germans and Austrians favored a "lone fighter," whereas the managers from Switzerland did not support this style.

Thus, the model is cross-culturally supported. The aforementioned studies demonstrate that the political and economic conditions should be considered as part of the situational factors. Some social situations give leaders more choices in using different decision-making strategies. As can be noticed in the case of women and the Polish managers before the fall of communism, the leader is devalued if he or she does not behave according to the expectations of the majority.

Different definitions of leadership that was developed mostly in North America and was validated in various countries or across men and women in leadership positions have been reviewed, and the influence of culture on leadership has been discussed. In summary, it seems that the validity of a universal characteristic of leaders is questionable. In contrast, contingency or situational leadership models showed more cross-cultural validity. However, culture must be considered as a situational factor, a moderator that influences the relationship between the leaders' traits, behavioral styles, and effectiveness.

3.3. Quasi-Emic and Emic Approaches: Implicit Leadership Theory

The previous sections reviewed the results of etic studies, in which the leadership theory, model, or measure was transported from one culture and applied to another to validate it. In these studies, however, implicit leadership theories of people of various cultures have not been the focus of research. It is through the emic approach that cultural uniqueness of leadership may come to light.

To illustrate the importance of subjective meaning of leadership in various cultures, the word that describes the activity in the workplace in different languages can be compared. For example, in English, the word "business" is used as a generic term for what goes on in the workplace. This word—denoting the condition of being busy—implies a sense of activity and energy to accomplish goals. In Spanish, however, the term used to refer to the same exchange is *negocio*. Etymologically, the root of this word is similar to negotiation, and it brings to mind exchange, social interaction, and give and take. Therefore, the comparisons of words used in different cultures that refer to activities in the work place may reflect the perspectives and priorities of these cultures. That is, in one culture the main focus may be getting the job done, in another, establishing relationships. Thus, cultural nuances are depicted by slight variations in the priorities on task versus relationships.

In cross-cultural leadership, two sets of emic studies can be identified: first, those that were inspired initially by Western conceptualizations and then arrived at an emic concept, which are referred to as quasi-emic; second, those studies that emerged empirically from within the culture, known as emic.

3.3.1. Quasi-Emic Approach

Three main leadership behavior models have emerged from the studies of leadership behavior in Iran, India, and Japan. These models did not support the two-factor model of leader behaviors (consideration/people orientation and instrumental/task orientation) that were present in North America. We refer to these endeavors as quasi-emic, because they were originally based on theories or measures that were inspired by researchers in North America. By examining the findings in each Asian country, similarities and uniqueness can be noticed.

In Iran, Ayman and Chemers reported that the behaviors describing consideration and initiating structure behavior style were both highly correlated with "My leader is like a father to me" and "My leader is effective"; the two styles were inseparable. Comparing the results of the studies in Iran, in which the leader was either American or Iranian, showed that Iranians maintained a perception that the two types of leader behaviors need to occur simultaneously for a leader to be effective and acceptable. Thus, the label "benevolent paternal" was coined for the Iranian leadership.

To illustrate the significant contrast of acceptable leadership behavior in Iran and the United States we can consider this anecdote: Iranians in the United States. have commented with amazement that their bosses are "all about business" and are impersonal, which is judged as inappropriate and rude. For example, it is said that first thing in the morning, the bosses walk into the offices, and without greeting employees and asking about their well-being, state their business, make task requests, and leave the room. However, when Iranians share this experience with their American peers, the response they get is that this is an acceptable behavior at work. That is, in the U.S. workplace, it is acceptable to forego exchange of social pleasantries. In Iranian culture, however, the expectation is for the leader to be personably warm and courteous as well as professional.

Examining the application of quasi-emic approach in India, J.B.P. Sinha and colleagues explored leader behaviors. They used the two leader behaviors of consideration and initiating structure, but also added culturally unique leadership behaviors they labeled as "nurturant-task." As a result, a unique conditional behavior that resembled the benevolent paternal leadership style in Iran emerged. In other words, leaders' behaviors included simultaneously expressed demand for getting the job done and caring for the person.

Concurrently, in Japan, Misumi developed a PM (performance/maintenance) theory of leadership behavior. He took the concept of performance-emphasis and maintenance-emphasis functions from Cartwright and Zanders' conception of leadership. The concepts resembled other leader behaviors popular in the United States, such as consideration and initiation of structure. However, PM is distinct from the leader behavior models in the United States. Misumi's theory emphasized that the two behaviors of attending to the task and to the people are not polar opposites, nor do they function separately. In addition, Misumi believed that leader behavior cannot be generic but should be put in context of the job. Therefore, not only was the PM model similar to the other Asian studies in the inseparableness of leader behaviors but it also mentioned the importance of context.

In summary, there is similarity across the studies in Asia, where the accepted leadership behavior style is a combination of consideration and initiating structure. In these cultures, focusing on only one aspect of work, that is, either getting the job done or attending to people, is not considered appropriate. The social protocols and pleasantries are significant elements of effective leadership. In these societies, business is about building trusting relationships through which the job gets done, not just about getting the job done.

In addition to the above studies, Schein and her colleagues conducted a series of studies around the world examining the image that people have of a manager compared to the image they have of a woman and a man. The results of these studies have shown that across cultures; there is a male image of a manager. This finding is very significant, as the number of women managers is increasing, and they will be faced with expectations that may be contrary to their style and habit. Interestingly, this may be the first universal finding in leadership that was not moderated by culture.

In summary, each culture cultivates a unique set of expectations of their leaders. The universal concept that emerges from these studies seems to underline the fact that leaders need to attend to task and to relationship. Across cultures, the importance of these two styles and their distinctions seem to vary. In Asia, it seems that effective leaders are expected to engage in task- and people-oriented behaviors simultaneously.

3.3.2. Emic Approach to Leadership

Earlier, the various types of cross-cultural approaches and the role of the endogenous and exogenous researcher were discussed. The previous section reviewed the results of the findings of studies conducted

by endogenous researcher. Based on their interpretations, an emic understanding of the leadership process in Asia evolved. This section presents two studies, in which, through a series of interviews and observations, two exogenous researchers explored leadership in China and Japan.

The first study's objective was to understand women leaders in China at the early stage of that country's decision to open itself to a capitalistic economy. Through a series of interviews, Korabik (1990) contended that Chinese women managers, compared to men, attributed their success mostly to interpersonal social skills. On the other hand, most of the men managers, as compared to women managers, attributed their success to skill, training, and experience. Interestingly, both Chinese men and women managers said, "It is possible to be both tough and caring. I try to combine patience and firmness." This finding reconfirmed the quasi-emic models of Asian leadership, in which the two task and interpersonal behaviors needed to occur simultaneously.

Pascale, in a 1978 case study of Japanese managers' style of leadership, presented comments from Japanese managers about how American and Japanese mangers differ in their style. As an example, he quoted a Japanese manager as saying, "You Americans are fond of announcing things. It sets everything astir." Pascale described the Western leaders as those who are "out standing," whereas Eastern culture values leaders who stand "in" rather than "out." The general but important difference between Japanese and American managers is that Americans are more explicit in their communication and the Japanese more implicit.

The emic studies of leadership, therefore, portray cultural subtleties in the leadership process. For example, across the world, leaders are expected to act more masculine. Also, if one were to use a continuum in which one end represented attention to performance and the other interpersonal relationship, the emic studies have demonstrated that Asian leaders seem to put more priority on a balance of interpersonal relationship and task accomplishment and see them as intertwined, whereas managers in the United States may prioritize task accomplishment at the possible expense of interpersonal relationship. The culture of men compared to women is typically more focused on task activities than emotions and interpersonal relationship. Hence, one could conclude that in the United States, men managers compared to women managers put more priority on performance than on interpersonal relationships. Also, one could surmise that since Asian managers seem to combine the relationship and task behaviors, their culture would be more similar to that of American women managers than to American men managers.

4. NEW TOPICS IN CULTURE AND LEADERSHIP

In recent years, three topics have emerged in the study of culture and leadership. One is related to the method of research in the study of cross-cultural leadership and the call for a multicultural team. Another is leadership when cultures collide, that is, when leaders operate in a multicultural environment. The third topic focuses on the evolving role of a transnational leader, that is, the difference between an international leader and a global leader.

4.1. Multicultural Research Teams

The majority of studies discussed so far have examined either the validity of one leadership paradigm cross-culturally or the unique meaning and function of leadership in a given country. Several recent research teams have explored culture and leadership. The GLOBE project, a recent cross-cultural research endeavor, has systematically examined similarities and differences of respondents over the question of leadership expectations and prototypes in several industries within each country.

In the late 1990s, House, Hanges, and Ruiz-Quitanilla described the GLOBE project as a multi-phase and multi-method research project that included collaborators representing various cultures around the world. They reported that the project presented data from 62 cultures and included 170 indigenous and exogenous social scientists. In their view, the concept of leadership is defined as "the ability of an individual to influence, motivate and enable others to contribute toward the effectiveness and success of the organizations of which they are members." The key research questions of this project focused on identifying both universal and culture-specific leadership attributes, behaviors, and organizational practices. Additionally, it identified in what way culture affected the type of leadership attempted and the effective leadership, and it presented the comparative standing of various cultures on the cultural dimensions. Finally, the project examined the potential for an underlying theory that could explain both national differences and the universal leadership attributes and practices.

To accomplish this, three measures were developed. One focused on the cultural and societal values attached to an effective leader (i.e., what they should be) and how they observed leaders behaving (i.e., what they are). The second measure addressed the practices of their organization. The third measure examined the attributes that impede highly effective leadership. (A detailed representation of the method and sampling of the GLOBE project is beyond the scope of this article.)

This project will eventually allow the examination of similarities and differences across cultures based on agreed-upon definition of leadership. Nonetheless, the preliminary findings attest to the same phenomenon that has been stated throughout this article. That is, although there may be some general universal qualities of leadership behavior, the acts manifesting them will vary and have different meanings across cultures.

4.2. Cultural Collision and Leadership

Multicultural work settings have been around for centuries. Immigrant workers and cross-national trade have existed around the globe for centuries. However, the intensity of cross-cultural exchanges has a more noticeable economic impact in today's global society than in the past. In previous centuries, those with resources and military might imposed their ways on others. Now, the competitive global economic market does not allow for a company or a nation to mistreat others for long, as there are competitors and potentially better options. Also, the presence of global communication systems has increased the accountability for social responsibility and human rights in business exchanges.

In today's global economy, expatriates can be emigrant workers or expatriate managers of large, multinational corporations. Leader–member exchange theory can assist understanding the impact of culture on leadership in a multicultural setting.

The leader–member exchange defines leadership as a relationship. It has demonstrated that leaders have varying levels of intensity in their relationships with different employees. When the leader and the subordinates are in a close, trusting relationship, they enjoy an enriched relationship in which the leader is more apt to engage in transformational behavior styles. When the subordinates are at the periphery of the team, are not closely involved with the leader, and do not have a strong trusting relationship, the leader tends to behave more transactionally with them. Leader–member exchange theory, therefore, demonstrates that like any relationship, the leadership process also goes through a maturation process, from being strangers to being confidants. What makes this model appealing to multicultural setting is that at the core it assumes that leaders and subordinates develop unique dyad relationships.

In a multicultural setting, it is conceivable that leaders will develop closer relationships with subordinates who are similar to them; the leader may stay distant from those who are different. The research on this model is still young, yet the opportunities to explore the role of culture in leader–subordinate dyad relationship building are promising. Exploring the cultural factors that contribute to building strong leadership relationships in multicultural settings is necessary.

Independent of this theory of leadership, practitioners and cultural anthropologists have developed cultural guides to help expatriate leaders adjust to working in cross-cultural settings and with people from various cultures. However, in this type of cultural awareness training, two issues are overlooked. The first is that the cultural prototypes are used. That is, in this training, the leader is told what typically may occur in a country or what to expect from people of a certain culture. This can create stereotypes, as most people do not fit the cultural prototypes or stereotypes, and so can lead to misunderstanding. Second, this method of cultural awareness assumes that culture is static. Trompenaars and Hampden-Truner stated that international managers are faced with layers of cultural demands that are fluid and need balancing. In addition, in this type of training there is a tendency for the leader to try to act similar to the natives. This can be a no-win situation, as it has been demonstrated that natives do not trust expatriate leaders who act like them. Consequently, the value of knowing about the norms of the other culture is the awareness of the range of accepted behaviors that can equip leaders with alternative explanations when they observe behaviors that may be strange to their own cultural norms.

In summary, our knowledge about leadership in multicultural settings is still evolving. For years, cross-cultural trainers such as Foa and Triandis have recommended that those interacting in cross-cultural settings defer judgment and view situations through the eyes of the member of the other culture. They also developed the cultural assimilator training programs, which teach the individual faced with a heterogeneous team to be aware of his or her cultural assumptions and potential biases, as well as the beliefs

and assumptions of the others. This technique trains the sojourner or expatriate to be more attentive and to think before acting.

Therefore, it is most helpful for the leaders in multicultural settings to be aware of the cultural norms of their own culture and those of the others. This awareness gives them a perspective of the range of behaviors to be expected. In addition, keeping an open mind and asking questions before judging is the best way to adjust to a team composed of various cultures and styles.

4.3. Global versus International Leadership

Many scholars use the terms global and international interchangeably. A number of scholars, however, use the terms to refer to two distinct mindsets. International leaders are those who have the tendency to impose their own cultural values and policies on others. On the other hand, global leaders perceive complexity and are inclusive, respecting people of all cultures equally, and are concerned about what is the best solution in the long run. This strategy is not much different than that referred to earlier as the collaborative approach to managing conflict.

Two forces affect the leaders who cross country boundaries or take part in cross-cultural functions. One is that most managers are faced with a universality of the technology and education in various fields. The other is that the international managers are faced with the diverse expectations and norms of the cultures they visit. The global leader has the ability and mindset to create the needed balance and harmony between these two. The international leader, though, will focus on the universal, and may not be skilled in or aware of the important cultural needs and expectations. The global leader incorporates this diversity when establishing his or her business plan and policies.

Global leadership requires awareness and acceptance of the diversity of the needs present in the world. In global leadership, the cultural interests and values of all are included; this process, due to its complexity, requires more time. Global leadership is inclusive and, through a balancing process, brings a sense of unity in diversity. As this is a new phenomenon, its representation is limited and difficult to examine empirically. But it is the type of leadership needed for the global society to progress and prosper. Some have proposed that perhaps a federation, committee, or council representing global-minded individuals with different cultural background is the best choice for a global leadership.

5. SUMMARY AND CONCLUSION

As this article has demonstrated, the study of culture and leadership is very complex. The progress in exploring this symbiotic relationship was explored through various perspectives, and the importance of including culture in understanding a social phenomenon such as leadership in a diverse and culturally heterogeneous setting was well documented. Imposed etic, emic, and derived etic studies were reviewed, and their findings were summarized. Overall, although the various approaches to leadership were validated in different cultural settings, these findings need to be considered with caution.

It seems that most countries, other than those with Anglo-Saxon background, place importance on the interpersonal relationship aspect of leadership. Also, the most powerful aspect of culture in determining leadership values or behaviors seemed to be economic environment such as urbanization. Additionally, although the various definitions of leadership functions were validated across cultures, the specific behavior that is perceived by one culture as contributing to one style is not the same across cultures. This fact alerts the sojourner managers of a multicultural team to be attentive as to how their specific acts may be interpreted differently by different members of their team.

In conclusion, to understand leadership in various cultures or globally, one must review multiple studies and approaches across time. This article provided a glimpse into the complexity present in the interplay between culture and leadership, highlighting the dangers of oversimplification and overgeneralization of the universality of leaders' behaviors. At the same time, some of the paradigms that have evolved so far in leadership, in principle, seem to be working globally. The challenge of a multicultural and global leader is to be attentive to the nuances present in different cultural groups and be accepting of them; in essence, to pursue a balance in unity and diversity.

See Also the Following Articles

Conformity across Cultures ■ Cross-Cultural Psychology, Overview ■ Industrial/Organizational Psychology across Cultures ■ Power, Authority, and Leadership ■ Training, Cross-Cultural

Further Reading

Ayman, R. (2004). Situational and contingency approaches to leadership. In J. Antonakis, A. T. Cianciolo, & R. J. Sternberg (Eds.), *The nature of leadership* (pp. 148–170). Thousand Oaks, CA: Sage.

Ayman, R. (2002). Contingency model of leadership effectiveness. In L. L. Neider, & C. A. Schriesheim (Eds.), *Leadership* (pp. 197–228). Greenwich, CT: Information Age.

Ayman, R., Kreicker, N. A., & Masztal, J. J. (1994). Defining global leadership in a business environment. *Consulting Psychology Journal: Practice and Research*, *46*, 64–77.

Chemers, M. M. (1997). *An integrative theory of leadership*. Mahwah, NJ: Lawrence Erlbaum Associates.

Chemers, M. M., & Ayman, R. (Eds.). (1993). *Leadership theory and research: Perspectives and directions*. New York: Academic Press.

Early, P. C., & Erez, M. (Eds.). (1997). *New perspectives on international industrial/organizational psychology*. San Francisco, CA: The New Lexington Press.

Gelfand, M. J., Raver, J. L., & Ehrhart, K. H. (2002). Methodological issues in cross-cultural organizational research. In S. G. Rogelberg (Ed.), *Handbook of research methods in industrial and organizational psychology. Blackwell handbook of research methods in psychology* (pp. 216–246). Malden, MA: Blackwell Publishers.

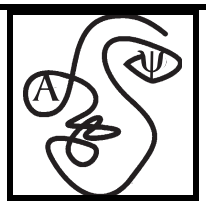
Geertz, C. (1973). *The interpretation of cultures*. New York: Basic Books.

Lewis, D. R. (1999). *When cultures collide: Managing successfully across cultures*. London: Nicholas Brealey.

McCall, Jr., M. W., & Hollenbeck, G. R. (2002). *Developing global executives*. Boston, MA: Harvard Business School.

Segall, M. H., Dasen, P. R., Berry, J. W., & Poortinga, Y. H. (1990). *Human behavior in global perspective*. New York: Pergamon.

Trompenaars, F., & Hampden-Turner, C. (1998). *Riding the waves of cultures: Understanding diversity in global business* (2nd ed.). New York: McGraw Hill.



Learning

Gary D. Phye
Iowa State University, Ames, Iowa, USA

1. Academic Learning
2. Learning as Enduring Change
3. Defining Elements
4. Types of Personal Knowledge Construction
5. Declarative Knowledge
6. The Role of Practice
7. Procedural Knowledge
8. Student Self-Efficacy: A Sense of Competency
9. Strategic Knowledge
10. Implications for Knowledge Assessment
11. Developmental and Individual Differences
12. Conclusion
    Further Reading

GLOSSARY

declarative knowledge Involves knowledge of facts, concepts, vocabulary, and the usual bits of information that are stored in memory.

formative assessment The observation or measurement of student learning outcomes prior to and during the learning experience. Formative evaluation provides a starting point for determining change as well as periodic assessment during the learning episode.

no child left behind (NCLB) Federal legislation that was enacted to redefine the manner in which federally supported programs are held accountable for the impact of educational interventions on all students in the classroom. Target behavior for defining impact is the improvement of academic achievement in mathematics, reading/literacy, and science.

pedagogy The art and science of teaching.

problem-solving transfer The ability of students to apply prior knowledge in the form of academic problem-solving strategies to new and novel problem situations in academic subject matter areas.

procedural knowledge This type of knowledge is observed in the classroom when students demonstrate the ability (knowing how) to combine, refine, incorporate, and accommodate (in the Piagetian sense) declarative knowledge so that it can be used in a course of action.

self-efficacy Personal beliefs concerning one's capabilities to organize and implement actions necessary to perform behaviors at designated levels (efficacy expectations).

self-regulation The process whereby students personally activate and sustain behaviors, cognitions, and affects, which are systematically oriented toward the attainment of learning goals.

strategic knowledge This type of knowledge is observed in the classroom when students' demonstrate "knowing when" to use prior knowledge that is in the form of declarative and procedural knowledge. This type of knowledge is observed in problem-solving situations.

summative assessment The assessment that occurs at the end of an educational intervention or teaching episode. A comparison of formative and summative assessment is the basis for the determination of change.

Academic learning is receiving increased attention in the early 21st century. The educational shibboleth is "no child left behind." This federal mandate has impacted the nation in a way that will continue to reverberate for several years. At the classroom level, teaching practices are being impacted by the emphasis on accountability

©2004 Elsevier Inc.
All rights reserved.

and impact measures that are being implemented. At the school district level, the accountability and impact data are to be used for policy analysis and to guide educational reform. The professional educator involved with the evaluation of educational practices and programs must produce credible data that meet the criteria of scientifically based evidence. In all cases, the target behavior for these efforts is the improvement of student achievement in reading, mathematics, and science. Assuming that we send our children to school to learn these academic subjects, we must also ask the questions, "what do we mean by academic learning and what are the processes by which students develop their knowledge of academic subjects?" These two questions serve as the focus of this article. Coming from the psychology side of educational psychology, the context is current research findings from the memory, cognition, and learning literature that have potential application to educational practices. Thus, it goes without saying that the perspective is "learner centered" and the unit of analysis includes both the individual and the group, as implied in the phrase "no child left behind."

1. ACADEMIC LEARNING

Learning as a proper noun has so many meanings that it is almost impossible to communicate effectively without first adding a qualifying term. For example, in the applied setting we call the classroom, there are different types of learning theories and models. Some theories describe and explain motor learning that serves as the basis for motor skill development. In other cases, some learning theories focus on learning that serves as the basis for social skills development. However, the focus of this article is academic learning that relies on cognitive learning theories and information processing models to describe and explain the development of personal knowledge about the subject matter we teach in the classroom.

I use the term academic as a descriptive adjective to communicate in one sense something about the nature of the thinking and remembering processes a student must engage in to successfully interact with classroom academic demands. In a second sense, I also draw attention to the fact that the curriculum content we commonly refer to as literature, mathematics, and science has a logically organized structure that must be understood by the student. In other words, in addition to simply acquiring information in the form of vocabulary, facts, or concepts, these building blocks for personal knowledge must be organized in a manner consistent with the epistemological demands of the subject matter being taught. Examples include knowledge of the syntactical structure of the language in which a third-grade student is writing, the logic for translating and representing a fifth-grade math story problem as an equation to be solved with the use of appropriate solution algorithms, and the formal prepositional logic involved with the ability to use combinations and permutations when solving ninth-grade algebra problems. From these examples, it is readily apparent that academic learning is quite different from social learning or motor learning, which also play a major role in students' efforts to adjust and cope with the demands of schooling.

2. LEARNING AS ENDURING CHANGE

Having identified academic learning in terms of a "type" of learning activity, a closer look at a commonly accepted definition of learning will serve as a means of considering what is frequently referred to as the learning process. In the third edition of *Learning Theories: An Educational Perspective*, Dale Schunk offers a suggestion. Although experts disagree about the precise nature of learning, the following is a general definition of learning that is consistent with a cognitive focus and captures the criteria most educational professionals consider central to learning: "Learning is an enduring change in behavior, or in the capacity to behave in a given fashion, which results from practice or other forms of experience" (Shuell, 1986, as cited in Schunk, 2000, p. 2).

There are three elements to Schuell's definition of learning that deserve further attention (change, endurance, and practice) because these elements help define the level of analysis involved in the present use of the term personal knowledge. This level of analysis is also driving the new "no child left behind" (NCLB) legislated mandates. These mandates emphasize the idea that all children can learn, and the key to success according to NCLB mandates is in the assessment and analysis of learning from a scientifically based perspective. Within this context of academic learning our discussion of the construction of personal knowledge can be critically examined.

3. DEFINING ELEMENTS

Behavioral change or change in the capacity for behavior is basic to any attempt to distinguish between information

processing and learning. Daily, each of us processes information through contact with our environment as well as information retrieved from memory in the form of ideas, recollections, etc. However, some of this information processing is at a level that does not initiate change in our behavior or our capacity to adapt. Academic learning involves the development of new behaviors or capacities, not simply the replication of previous habits. Furthermore, the demonstration of change is not left to conjecture. Rather, change can only be demonstrated by measuring the behavior at two different points in time. Typically, this is prior to and following instruction.

Enduring change must have some permanency about it. In other words, the important elements of instruction to which a student is exposed should be remembered. In fact, our curriculums in the public schools are designed around this primary assumption. Unfortunately, the empirical evidence for enduring change is scanty at best. In our discussion of personal knowledge, we focus on the individual learner as the unit of analysis, not the class as a whole. This is in fact the reason that the NCLB mandates are framed within a consideration of the individual student (no child left behind). This in no way ignores the fact that academic learning occurs in a group context. Rather, it focuses attention on both the individual and the collective we call a classroom. This makes common sense when one considers that the basis for knowledge permanency is the individual student's memory system.

Practice is an interesting concept in the academic community. A basketball coach teaching motor skills will tell you that "practice makes perfect," a music teacher will insist that students practice between lessons, and the school administration will stress the practice of civility and social behavior. However, the classroom teacher responsible for academic content frequently views practice as a "hit-or-miss" proposition that gets attention during seatwork or homework. Interestingly, although all these practice episodes can be observed in any school, the episode that typically lacks any organization or rationale is seatwork or homework.

Together, these three basic elements of our definition of academic learning stress the idea that academic learning is not a one-shot instructional experience. It is a process of acquiring new information, the refinement and organization of what is already known, and the successful use of that knowledge. Academic learning is the product of practice that provides the basis for relatively long-term change in one's personal knowledge. Ideally, this change increases students' ability to be successful as they move from grade to grade where the curriculum requires more complex and specialized forms of personal knowledge.

4. TYPES OF PERSONAL KNOWLEDGE CONSTRUCTION

The following discussion is based on a common theoretical assumption that pervades educational psychology: Cognitive change must occur as a student progresses through the curriculum in order to be successful. Jerome Bruner, in his discussion of the spiral curriculum, made this point. Still valid today, this assumption is frequently presented in the form of a hierarchical schema, from simple to the complex, in terms of both the cognitive processing demands made on students and the nature of the subject matter to be learned. The following discussion takes this form and reflects a prevailing attitude among educational psychologists that the construction of personal knowledge can be identified as falling within three major typologies: declarative knowledge (concepts and facts), procedural knowledge (strategies, procedures, and schemata), and strategic knowledge (expertise in academic problem solving).

The study of knowledge can take many forms. Within the context of academic learning, the following analysis is functional in nature and involves an information processing perspective. The functional nature of the following discussion involves addressing the "what" and "how" questions of personal knowledge construction. Phye (1997a) stated,

> A particular value of the information processing perspective for the classroom teacher is the assumption that, while each student in the classroom will have a unique set of prior experiences that influences entry learning level, rate of learning, style of learning, and so forth, all nondisabled students in the classroom will be approaching the learning assignment with essentially the same set of processing components. Consequently, when learning is not successful, an assessment can be quickly conducted in order to determine cognitive processing strengths and weaknesses. (p. 53)

5. DECLARATIVE KNOWLEDGE

Basically, declarative knowledge involves those things teachers refer to as facts, vocabulary, concepts, etc. These bits and pieces of information are the building blocks for academic learning. These are the "basics" that

cut across all curriculum content areas. For the student, the construction process involves not only attending and remembering but also the bringing of meaning to this newly acquired information. By connecting new information with prior knowledge or using new information to refine what is already known, meaning is changing from the concrete to the complex.

Three examples make the point. Outside the classroom as well as in the primary grades, a commonly accepted mathematical fact is that $2 \times 2 = 4$. However, as one progresses through the mathematics curriculum, a part of any student's academic learning involves the following refinement of prior knowledge. The $2 \times 2 = 4$ mathematical fact applies only in a base 10 mathematical system. Thus, $2 = 2 = 4$ may not be a fact in other mathematical systems. This mathematical fact is refined or deemed to be "true" only within a specific context. For most students, this refined knowledge would have little or no relevance outside the classroom.

The second example is a common classroom learning experience that is observed when dealing with concepts in science. During the early grades, the concept of "animal" is introduced and students typically use the term to mean a large group of "nonpeople living things" of one type or another. However, as science lessons progress, the student is challenged to refine this general concept into a more specific organizational system that includes people within the subcategory of mammals and that mammals are still animals. This bit of declarative knowledge refinement has both high academic relevance and relevance outside the classroom.

Having touched on examples from mathematics and science, the last example comes from early reading. As a child progresses from kindergarten to the fourth grade, vocabulary building, fluency in reading, phonetic awareness, word attack skills, and comprehension are reading skills that are practiced in oral reading groups. As these basic reading skills become integrated, a child's reading performance improves. A primary goal of learning to read involves the ability to comprehend what is being read. One part of this process is taking individual words and their meanings and putting them together into sentences. Interestingly, in English many words have different meanings depending on whether they are used as a noun or a verb in a sentence. The word "desert" used as a noun means a dry, barren, often sandy region that can naturally support little or no vegetation. Desert used as a verb means to forsake or leave, especially when most needed. Obviously, the key to understanding the meaning of the word desert depends on having both meanings available and selecting the appropriate meaning based on the context (rest of the sentence).

Students process information every waking moment of their lives. Some of the information is organized and some is fragmented or organized around real-world survival skills. However, academic knowledge is a special case in which the organizational principles are unique to the subject matter being taught. Consequently, remembering becomes a valuable cognitive skill for any student. Thus, as we go back to our basic definition of academic learning, the emphasis on enduring change becomes more apparent. As learners, we remember facts and concepts not only so we can demonstrate the acquisition of new declarative knowledge but also so this new knowledge becomes prior knowledge to be added to and refined at a later point in time. The reading example described previously also demonstrated the vital role that memory plays in academic learning. The student who encounters the word desert must remember two different meanings and successfully retrieve from memory the appropriate meaning, depending on the cues provided by the rest of the sentence. The ability to do this fluently requires a great deal of practice so that the process of remembering meanings occurs almost automatically.

Reading fluency is a good example of an academic behavior that improves with practice and is a critical academic survival skill. By the intermediate grades, students are asked to take increasingly more responsibility for their own learning. Most classroom environments change at the middle school level and students are expected to acquire a great deal of their own declarative knowledge by reading texts, working through Web-based instructional modules, etc. This expectation on the part of educators suggests that students should be practicing their reading skills during the primary grades so that they can successfully adapt to the demands of the middle school.

6. THE ROLE OF PRACTICE

Although practice is an essential element of the learning process, a frequently asked question is, "Does practice make perfect?" It goes without saying that practice of some type is typically more beneficial than no practice. However, not all practice is equally effective. This statement is simply the recognition that practice plays an important role in two quite different ways. In terms of acquiring academic knowledge, repeated exposures or repetition provide the basis for memory storage that is critical for remembering. In terms of reading

vocabulary, the association of the word desert with its meaning and the remembering of this pairing will require several repetitions in order for the information to be stored in long-term memory. This may be accomplished with a simple practice strategy that involves developing a set of what were called "flash cards" when I was a youngster. However, to successfully retrieve and use the appropriate meaning across a number of different reading passages requires the successful application of prior knowledge. The ability to successfully apply one's knowledge is really a disposition or personality characteristic. This disposition is also acquired through practice and has been identified by Albert Bandura as a sense of "self-efficacy" or competence. In other words, successful practice strategies involve both multiple exposures to the task and actively seeking potential applications of prior knowledge. This is a simple example of "training for transfer."

7. PROCEDURAL KNOWLEDGE

Within the context of academic learning, procedural knowledge is demonstrated when a student can combine, organize, refine, or accommodate declarative knowledge so that a course of action can be taken. This is "knowing how" to use declarative knowledge (what one knows). Procedural knowledge in the classroom is what we commonly strive to teach and we ask our students to learn.

In teaching circles, constructivist or student-centered approaches to instruction place a great deal of emphasis on promoting procedural knowledge. This is frequently the case in the areas of science education and mathematics education. The caveat is that students must have a solid background of declarative knowledge in order to successfully move to a more complex level of cognitive processing. The cognitive processing that serves as the basis for procedural knowledge is domain specific. For example, if we think of algebra as a part of the mathematics domain, there are basic mathematical facts and concepts (declarative knowledge) that have been learned in earlier grades that serves as the foundation for success in algebra. However, in algebra we are also asked to take our prior knowledge about mathematics and use that knowledge to develop new rules, concepts, strategies, algorithms, etc. This is really a case of taking what we know and learning how to use that knowledge to develop new and more complex ways of thinking.

For example, by the time students reach middle school, where they usually take algebra, the teaching of reading shifts from a focus on teaching basic reading skills to using reading skills to promote self-directed learning in literacy. At this point in students' academic careers, emphasis is on procedural knowledge involving "knowing how to read" in order to successfully perform other educational tasks (write a poem, give a persuasive speech, read a science or mathematics textbook, etc.). These complex academic tasks are built on the declarative knowledge any student remembers from the earlier grades. This is why remembering (memory storage and retrieval) is the key to academic learning because complex subject matter that students are asked to learn is always built of declarative knowledge and simpler forms of procedural knowledge.

These descriptions of our expectations for students as they progress through the primary and middle school grades again draw our attention to the key academic learning element of enduring change that is couched in practice (seatwork, homework, study hall, etc.). Interestingly, there is little evidence in the research literature that teachers systematically teach study skills. This means that teachers should not only present information in the classroom but also teach students how to use the information being conveyed. Thus, teachers must teach not only declarative knowledge (what) but also procedural knowledge (how). A position frequently taken by educational psychologists is that domain-specific procedural knowledge strategies must be a part of any teacher's instructional repertoire.

8. STUDENT SELF-EFFICACY: A SENSE OF COMPETENCY

Self-regulation or self-regulated learning emphasizes the motivational side of human behavior. As such, it is a complementary view of the information processing approach to academic learning. Metaphorically, cognitive theory and motivational theory are two sides of the "coin" called academic learning. Research efforts by educational psychologists studying self-regulation emphasize how learners direct their thoughts, feelings, and actions toward the attainment of their academic goals.

Bandura's concept of self-efficacy was discussed previously within the context of study or practice. Generally speaking, self-efficacy is the exercise of control by the individual in the daily task of adapting to a complex environment. Self-regulation and the development of a sense of self-efficacy grow out of a complex interaction of behaviors, interpersonal factors,

and environmental demands. In the classroom, a simplified example involves teacher variables, curriculum demands, and students' sense of success with the curriculum being taught (math, reading, etc.). This sense of success translates on a personal level for the student as being competent in the subject being taught. Furthermore, the success breeds a sense of self-efficacy: "I can do this, I understand how to take what I already know and use it to learn more about this because I understand what it takes to get better." "I am a good reader." "I am good at math."

Self-efficacy was introduced earlier within the context of practice because of the position taken by Bandura that it is the demonstration of being competent (a good reader or good at math) that precedes the sense of self-efficacy. Over the long term (from kindergarten to grade 12), it is the sense of self-efficacy in academic subject matter areas that provides the motivational basis for successfully making the transition between procedural and strategic knowledge construction by a student.

This sense of self-efficacy is what sustains a student when academic problem solving becomes complex and difficult. When confronted with a complex academic problem, a successful problem solver has sufficient declarative knowledge to bring to the task; sufficient procedural knowledge to synthesize, integrate, and evaluate the declarative knowledge; and can strategically (knowing when) translate this into a successful solution for a specific problem. It is self-efficacy, the sense of being able to exercise personal control over the process of knowledge construction in an academic setting, that characterizes strategic knowledge.

9. STRATEGIC KNOWLEDGE

Strategic knowledge is viewed as an extension of declarative and procedural knowledge. Whereas declarative knowledge involves learning facts, concepts, etc. and procedural knowledge involves learning how to use declarative knowledge to construct domain-specific rules, algorithms, etc., strategic knowledge involves knowing when to use prior knowledge. In other words, when confronted with a novel or problem-solving situation, emphasis is on the retrieval of prior declarative and procedural knowledge that has been stored from long-term memory. When this type of information processing is successful, we observe academic learning behaviors such as problem solving in mathematics, critical thinking in science, or creative thinking in the writing of fiction. This last type of knowledge demonstrated as a learning outcome emphasizes the active construction of a learning outcome and places the student at the center of the teaching–learning process. In addition to reflecting a competency dimension, primary consideration is given to motivational aspects of academic learning. This involves helping students develop a disposition frequently referred to as intellectual curiosity or academic expertise.

The development of strategic knowledge may be based largely on student characteristics that are very difficult to teach. Furthermore, some students may simply have difficulty with this level of cognitive functioning because of their cognitive limitations or personality characteristics.

Regardless, a teacher should provide opportunities that foster the development of strategic knowledge for those students who are capable. Strategic knowledge is a primary learner outcome of a self-regulated learning process we frequently identify as academic problem solving. Academic problem solving is observed when students can take primary responsibility for demonstrating the learning outcomes of problem identification, problem representation, strategy construction, and solution verification (evaluation).

Emphasis is on the student taking responsibility for identifying a problem, representing the problem, developing a strategy for solving the problem, and then evaluating the success of the strategy employed. This academic problem-solving perspective has been a major contribution to the teaching literature that is based to a large extent on the research efforts of Richard Mayer. These stages of the problem-solving process can be taught in any of the previously identified academic areas (mathematics, science, and literature). Any self-regulated learning process places a premium on remembering—the use of prior knowledge. The student is responsible for knowing what, how, and when. As demonstrated by Barry Zimmerman and colleagues, if we are successful, some of our students will become self-regulated learners. The choice is theirs, and the reward will be ours.

10. IMPLICATIONS FOR KNOWLEDGE ASSESSMENT

The implications for pedagogical practice are again derived from our definition of academic learning. Consistent with the NCLB legislation, the focus is on every child. This calls for the recognition of individual

differences in the classroom. In a single classroom on any given day, some students will be dealing with the development of declarative knowledge in a subject area while others will be grappling with issues involving procedural knowledge. A few may even be working at the strategic knowledge level of construction. On a daily basis, this means that classroom teachers will have to be well grounded in both the subject matter being taught and effective teaching strategies. In addition, since change is the focus of our definition of academic learning, this means that when appropriate, teachers are responsible for guiding students from one level of knowledge construction to another. This is a long-term view and involves change over the academic year within a grade. Consequently, for future change to occur, the student must remember today's lesson, which is tomorrow's prior knowledge. The implications for practice are obvious since the effective use of memory storage and retrieval is not easily accomplished.

Formative and summative assessment of academic learning is required since enduring change can be verified only to the extent that it can be measured. As mentioned previously, measurement at a minimum of two points in time is required for the assessment of change. Furthermore, a classroom environment in which student progress is monitored effectively would require multiple assessment techniques. The major obstacle facing most teachers is the assessment issue. There are numerous sources available for teachers that provide the basis for developing assessment activities that measure declarative knowledge. Most of these activities involve the assessment of memory retention using standard teacher-made tests.

The assessment of procedural knowledge and strategic knowledge also requires the assessment of memory retention, but it must also include the assessment of a student's ability to use the knowledge in a manner that reflects the transfer of higher order thinking skills (e.g., Bloom's taxonomy). This means that the assessment will require evidence of strategy transfer or problem-solving transfer on the part of students. Thus, as we promote the development of procedural and strategic knowledge through the use of instructional practices that foster such thinking, we must not forget that assessment practices must align with the instructional efforts and the learning expectations we have for our students. In simple terms, there are two assessment questions we must ask ourselves as teachers: Do our students remember what we taught, and can they effectively use what they remember?

11. DEVELOPMENTAL AND INDIVIDUAL DIFFERENCES

A model for knowledge assessment is provided in Table I. As mentioned previously, declarative knowledge can be assessed in the traditional manner and is frequently a one-time occurrence. In contrast, procedural knowledge and strategic knowledge require both a measure of memory retention and a measure of transfer. This dual-assessment procedure is required because transfer is based on prior knowledge. If we simply measure transfer performance and transfer performance is poor or nonexistent, as teachers we cannot determine if our students fail because they cannot remember or because the cannot use their declarative knowledge. In other words, memory retention is a necessary but not sufficient condition for transfer. Thus, teaching for transfer is based on teaching for declarative knowledge but also involves helping students understand how and when to use their knowledge.

Having made a distinction between an assessment procedure for declarative knowledge and procedural or strategic knowledge in terms of single or dual assessment, a distinction is also to be made between the assessment of procedural and strategic knowledge. In contrast to procedural knowledge, in the discussion of strategic knowledge, the volitional nature of strategic knowledge was emphasized. During procedural

TABLE I
Model for Functional Knowledge Assessment

Type of learning outcome	*Memory retention performance*	*Transfer performance*	*Kind of knowledge*
Nonlearner	Poor	Poor	None
Non-problem solver	Good	Poor	Declarative
Guided problem solver	Good	Good	Procedural
Self-directed problem solver	Good	Good	Strategic

knowledge assessment, if students can transfer procedural knowledge "when asked to do so," guided understanding is demonstrated. From a problem-solving transfer perspective, these students are demonstrating the last three stages (problem representation, strategy selection, and solution verification) of the aforementioned problem-solving process. Unfortunately, we lack information about problem identification. The assessment of strategic knowledge, where no teacher guidance is provided, provides information about students' ability to identify problems. Here, students must volitionally construct an awareness of what knowledge is appropriate, how that knowledge can be organized into a solution strategy, and when the strategies are to be employed. Such students not only demonstrate the learning outcome of understanding but also the characteristics of successful problem solvers. This distinction between the assessment of procedural and strategic knowledge must always be considered when developing performance assessments in the classroom.

12. CONCLUSION

The multiple assessment approach to promoting understanding and problem solving in the classroom is at the heart of the NCLB legislation. No child should be left behind as a nonlearner. Every child can learn to a level that is determined by the individual differences in the academic abilities each of us exhibit. The hope is that as teachers, we facilitate to the best of our ability the personal knowledge construction efforts of all our students.

See Also the Following Articles

Educational Achievement and Culture ■ Gender and Education ■ Intelligence in Humans ■ Learning Disabilities ■ Learning Styles and Approaches to Studying ■ Transfer of Learning

Further Reading

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Bruner, J. S. (1966). *Toward a theory of instruction*. New York: Norton.

Haskell, R. E. (2000). *Transfer of learning: Cognition, instruction, and reasoning*. San Diego: Academic Press.

Mayer, R. E. (2002). *The promise of educational psychology*. *Vol. 2: Teaching for meaningful learning*. Upper Saddle River, NJ: Merrill/Prentice Hall.

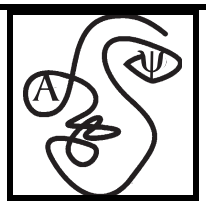
Phye, G. D. (Ed.). (1997a). *Handbook of academic learning: Construction of knowledge*. San Diego: Academic Press.

Phye, G. D. (Ed.). (1977b). *Handbook of classroom assessment: Learning, adjustment, and achievement*. San Diego: Academic Press.

Schunk, D. H. (2000). *Learning theories: An educational perspective* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Shuell, T. J. (1986). Cognitive conceptions of learning. *Review of Educational Research*, *56*, 411–436.

Zimmerman, B. J. (1990). Self-regulated academic learning and achievement: The emergence of a social cognitive perspective. *Educational Psychology Review*, *2*, 173–201.



Learning Disabilities

William D. Tilly III
Heartland Area Education Agency 11, Johnston, Iowa, USA

1. History of Learning Disabilities
2. Definition
3. Treatments for Learning Disabilities
Further Reading

GLOSSARY

formative evaluation The process of using data collected frequently and repeatedly over time to determine the effectiveness of a treatment program.

learning disability As used in this article, a heterogeneous group of disabilities that result in deficits in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, that may manifest itself as an imperfect ability to listen, think, speak, read, write, spell, or do mathematical calculation.

modality instruction The practice of assessing certain characteristics of the learner (i.e., modalities) such as whether the individual is a visual, auditory, or kinesthetic learner, and then matching treatments to the individual's strengths.

perceptual-motor training The most frequent form of "process training" in special education, . it assumes that children with learning disabilities, or at least some of them, have problems in sensory integration to which some or all of their learning problems can be attributed. The purpose of the training is to remediate these deficits.

response to intervention (RTI) An approach to assessment and treatment of persons with learning problems that can, in many cases, assist in identifying individuals with learning disabilities. In this method, students with specific learning problems are placed in a validated intervention programs and their "response to the intervention" is observed and used in decision making.

This article describes the development of our understanding of learning disabilities in the United States. It begins with a brief history of learning disabilities, then the struggles for a universal definition are described, along with the federal legal definition and the emerging consensus on definition within the field of learning disabilities. Finally, a sampling of both proven effective and proven less-effective treatments for learning disabilities is presented.

1. HISTORY OF LEARNING DISABILITIES

The existence of learning disabilities has been documented for approximately 200 years. The origins of what are currently called learning disabilities began with early attempts to match functioning within certain areas of the brain to human behavior. Early physicians noticed that patients with certain types of brain injury experienced specific types of behavioral problems. Most relevant to learning disabilities were patients with injuries to the left hemispheres of their brains, who in many cases experienced problems such as slow, laborious speech and the inability to name objects or persons. In the late 1800s, a more specific set of cases was identified in which persons' ability to speak and write remained

©2004 Elsevier Inc.
All rights reserved.

intact, but their ability to comprehend spoken or written words was impaired. These conditions were termed word deafness and word blindness.

From these studies, during the early 20th century, physicians began to notice children and adults who experienced what was at the time considered to be defects in their ability to recognize words and letters, despite no obvious brain damage or injury. The inference was made that perhaps these individuals suffered from some sort of neurological disorder that mimicked the effects of people with brain damage. Moreover, a number of cases appeared to run in families, providing further evidence for a possible physical cause. The medical diagnosis for this condition was congenital word blindness.

Beginning around 1920, and building on this early work in medicine and neurology, a group of American pioneers including Samuel Orton, Grace Fernald, Marion Monroe, and Samuel Kirk began landmark work on understanding reading disabilities. While still clearly linked to foundations in medicine, reading disabilities were attacked by these researchers and clinicians as much from an educational perspective as from a clinical one. Indeed, some of these pioneer's names have become synonymous with specific approaches to educational treatment for students with learning disabilities. For example, the Orton-Gillingham approach to treatment is a phonics-based, multisensory approach to instruction that uses the kinesthetic, auditory, and visual modalities for teaching decoding and spelling.

Parallel to developments in educational understanding of students with these disorders, additional work was being pursued between about 1920 and 1960 on the perceptual, perceptual-motor, and attentional abilities that often accompanied damage, injury, or dysfunction in persons' brains. Physicians and psychologists such as Kurt Goldstein, Heinz Werner, Newell Kephart, and William Cruikshank conducted many research studies documenting deficits such as hyperactivity, perseveration, figure-ground confusion, and distractibility. Again, it was reasoned that if persons with documented damage to their brains exhibited perceptual and attentional problems, persons with presumed damage to these brain areas might be expected to have these problems as well.

It was between approximately 1960 and 1975 when the field of learning disabilities began emerging as a discipline unto itself. On April 6, 1963, another pioneer in the field of learning disabilities, Samuel Kirk, told a parent advocacy group for children with "perceptual handicaps" that "Recently, I have used the term 'learning disability' to describe a group of children who have disorders in development, in language, speech, reading and associated communication skills needed for social interaction. In this group, I do not include children who have sensory handicaps such as blindness or deafness, because we have methods of managing and training the deaf and blind. I also exclude from this group children who have generalized mental retardation" (Kirk, 1975, p. 9). Most educators attribute coinage of the term "learning disability" to Kirk. Moreover, some of the parents who heard Kirk's speech in 1963 established the Association for Children with Learning Disabilities, which has now evolved into the Learning Disabilities Association of America. This organization is widely known as one of the most influential learning disabilities advocacy organization in the United States.

2. DEFINITION

2.1. Historical Definitions

One of the biggest challenges in the field has been coming up with a definition of learning disabilities that a large majority of people can agree with. Part of the challenge has been the heterogeneous nature of problems associated with learning disabilities. That is, persons with learning disabilities do not all have the same problems, and the types of problems experienced are very broad. While most persons with learning disabilities have problems with reading and the use of language (roughly 80%), not all do. Some have problems with memory, others do not. Still others have perceptual difficulties. And the list goes on.

In the years between 1960 and about 1975, there was considerable disorganization and disagreement on the specific nature of what constitutes a learning disability. Though Sam Kirk and his student Barbara Bateman had offered a series of definitions in the early and middle 1960s, definitions created by different advocacy and professional groups continued to stress different components of the disability. It was at this time that the federal government began taking interest in learning disabilities and began working toward definitions of its own.

2.2. United States Legal Definition

A series of federal task forces on learning disabilities and legislative actions occurred throughout the 1960s and into the early 1970s. These set the stage for the federal government's enactment of Public Law 94-142, the Education of the Handicapped Act, in 1975 (EHA). Coming on the heels

of the civil rights movement, EHA provided a mechanism for the federal government to provide funding to states to support the provision of special education services to students with disabilities. Based on the law, children with disabilities could not be excluded from receiving an education solely on the basis of their disability. Major components of the law included ensuring students with disabilities a right to a free, appropriate public education, heightened access to due process and procedural safeguards to ensure appropriate provision of services, and provision of an Individual Education Program that was reasonably calculated to provide benefit to students with disabilities (in this case, learning disabilities).

EHA was an impressive and ambitious step by the federal government. At the time, however, the promise of EHA could not be realized unless an acceptable definition of learning disabilities could be written. Recall that at the time, there was considerable professional disagreement on the nature of learning disabilities. Congress and the U.S. Office of Education considered much testimony, research, and advocacy in selecting the definition of learning disabilities that finally made its way into federal law. Indeed, many components of historical definitions were considered, and the final definition of learning disabilities was finally written. According to federal law, a learning disability is as follows:

> General. The term learning disability refers to a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, that may manifest itself as an imperfect ability to listen, think, speak, read, write, spell, or do mathematical calculation, including conditions such as perceptual disabilities, brain injury, minimal brain dysfunction, dyslexia, and developmental aphasia. Disorders not included: The term does not include learning problems that are primarily the result of visual, hearing, or motor disabilities or mental retardation, of emotional disturbance, or of environmental cultural or economic disadvantage [U.S.C. 300.7(C)(10)].

In addition to the definition provided in the U.S. federal law, the U.S. Department of Education provided regulations that further direct states and territories in determining the presence of specific learning disabilities. As stated in the regulations:

a. A team may determine that a child has a specific learning disability if:
   1. The child does not achieve commensurate with his or her age and ability levels in one or more of the areas listed in paragraph a.2 of this section, if provided with learning experience appropriate for the child's age and ability levels; and
   2. The team finds that a child has a severe discrepancy between achievement and intellectual ability in one or more of the following areas:
      i. Oral expression
      ii. Listening comprehension
      iii. Written expression
      iv. Basic reading skill
      v. Reading comprehension
      vi. Mathematics calculation
      vii. Mathematics reasoning

b. The team may not identify a child as having a specific learning disability if the severe discrepancy between ability and achievement is primarily the result of:
   1. A visual, hearing, or motor impairment
   2. Mental retardation
   3. Emotional disturbance
   4. Environmental, cultural, or economic disadvantage

Based on the language in the law and the language in the regulations, each of the 50 states and U.S. territories needed to establish specific procedures that would be used to identify individuals with specific learning disabilities.

2.3. Converging Professional Perspectives on Definition

The federal definition of learning disabilities set the stage for greater stability within the learning disabilities professional community and general movement toward agreement on a definition. One reason for this progress between about 1975 and 1985 was the considerable research conducted across the country that was funded by the U.S. Office of Education. Five learning disability research institutes were funded at major research universities across the country, each focusing on a different component related to understanding and treating individuals with specific learning disabilities in the schools. The institutes included the following:

- The Columbia University Institute, directed by Dale Bryant, focused primarily on information-processing problems of students with disabilities. Within this program, memory skills, arithmetic, study skills, basic reading, and reading comprehension (among other areas) were studied.
- The University of Illinois at Chicago Institute, directed by Tanis Bryan, focused primarily on the social competence of children with learning disabilities and their attributions of success or failure.
- The University of Kansas Institute, directed by Donald Deshler, focused primarily on school-based

interventions for secondary (adolescent) students with learning disabilities.

- The University of Minnesota Institute, directed by James Ysseldyke, focused their energies on the decision-making process related to identification of students with learning disabilities as well as methods for monitoring academic progress for students with learning disabilities. It was from the Minnesota Institute that curriculum-based assessment and curriculum-based measurement technologies were developed initially.
- The University of Virginia Institute, directed by Daniel Hallahan, focused their research efforts on students with learning disabilities who also had attentional problems. Additionally, the Virginia site also conducted research on the use of academic strategy training to improve the performance of students with learning disabilities.

During the 10 years following the work of the learning disabilities institutes, a series of new definitions of learning disabilities was offered by professional organizations and federal task forces. These definitions, though having some different nuances, all describe learning disabilities as (1) intrinsic to the individual, of presumed neurological dysfunction, (2) related to the use of language such as reading, writing, and reasoning, and (3) able to impact the individual in other areas of their lives such as social skills, mathematics, and attention.

Throughout the years between about 1985 and 2000, much additional research was conducted on the nature of learning disabilities. While there are many significant and important findings in this literature base, it appears that there are two large strands of research emerging related to the definition of learning disabilities. The first major area is in phonological processing. In 2000, a group of distinguished researchers (called the National Reading Panel) came together to examine and synthesize what is known about the acquisition of reading. In their report, these researchers synthesized the major skills that a person must have to be able to read proficiently. These include the following:

Phonemic awareness: The ability to hear and manipulate the sounds in spoken words and the understanding that spoken words and syllables are made up of sequences of speech sounds.

Phonics: Includes both the understanding that words are composed of letters that represent sounds and the ability to use letter–sound correspondence to pronounce an unknown printed word.

Vocabulary: Includes both expressive vocabulary, which is the ability of a speaker or writer to produce a specific label for a particular meaning, and receptive vocabulary, which requires a reader to associate a specific meaning with a given word (either read or spoken).

Fluency: The ability to read words with no noticeable cognitive or mental effort. It is having mastered word recognition skills to the point of overlearning.

Comprehension: This has been defined many ways by many authors. Comprehension includes active and intentional thinking in which meaning is constructed through interactions between the text and the reader.

The importance of these five "big ideas" in reading is that we now have both a roadmap for directly measuring the subcomponents of reading and specific strategies, linked with each of these ideas, that can be used to directly teach the skill. As noted before, a majority of students with learning disabilities have problems with reading, and this new knowledge provides us both a very specific way to identify specific skill deficits and a sensitive way to identify students with problems. As research progresses, we can expect this strand of research to provide increasingly effective treatments for students with learning disabilities.

The second major strand of learning disabilities research has to do with the identification of a biological basis for learning disabilities. Since the early years of learning disabilities identification, there has been a presumption of a neurological basis for these disabilities. Not until the late 1980s and 1990s has our technology caught up with our intuition and allowed us to study and document these findings. There are a number of different converging sets of data at this point. We know now from postmortem studies that the brains of persons with reading disabilities differ from those of persons without reading disabilities. We know from neuroimaging studies that there are differences in critical brain functions between proficient readers and nonproficient readers. We know also through systematic evaluation of heredity that about 40% of first-degree relatives of students with reading disabilities, speech and language disabilities, and spelling disorders also experience these problems.

2.4. Proposed Future Definitions

While all of these new research studies and findings were being created, the federal legal definition of learning disabilities (1975) has remained virtually unchanged. Our new research base, coupled with the practical and political realities associated with learning disabilities, has created a context for reexamination of the federal definition. A major political and practical

challenge is that since the inception of the Education of the Handicapped Act, both the number and percentage of students identified with learning disabilities in U.S. schools has consistently and steadily risen. This rise creates both fiscal and practical challenges. Currently, approximately 50% of all children with educational disabilities served under the IDEA are students with learning disabilities. There are many reasons for this rise, including that (1) the assessment and identification process in schools has been in many cases unreliable, (2) the use of technically inadequate tests is widespread, and often little functional information about how to teach a child with a learning disability is provided from the comprehensive evaluations required by federal law, and (3) different states use very different practices and criteria for determining the presence of learning disabilities. Moreover, some experts contend that learning disability numbers are increasing due to an increasing number of students who are what has been called curriculum casualties. That is, they have not learned because either the curriculum is not geared to meeting their learning needs or their teachers are not skilled in providing instruction in reading, language arts, and/or mathematics. When curriculum and instruction fail, special education is there to fill the gap. In short, the identification and special education service delivery process worked somewhat well from an administrative perspective of allocating resources, but it did not work well from the standpoint of remediating critical skill deficits nor in creating a consistent identification system for learning disabilities.

Added to our research base and the practical challenges we have experienced, additional evidence has been added in very recent years to our understanding of learning disabilities. Headed up by Reid Lyon, Chief of the Child Development and Behavior Branch of the National Institute of Child Health and Human Development of the National Institutes of Health, and a national team of researchers from many major universities, critical insights and definitive studies have begun to emerge that are causing researchers, policymakers, and practitioners to reconsider the way we identify and serve individuals with learning disabilities in our schools. Perhaps the most critical addition to our common knowledge base is that the prevailing practices used in our schools to identify and serve students with learning disabilities are problematic from the standpoint of improved outcomes for students. Our emerging knowledge is compelling.

1. We know that most students with learning disabilities have important problems with the use of language and acquisition of reading skills (approximately 80%) and that these problems begin early in life.

2. We know that in American society there are serious, important, and life-long consequences of not learning to read.

3. We have learned that there is a critical period for children to develop language and early reading skills. There are five major component skills that proficient readers master, typically in kindergarten to grade 3. They are phonemic awareness, phonics, vocabulary, fluency, and reading comprehension. All of these skills are directly teachable.

4. We know that the most effective interventions for reading problems are those that begin early and teach students in these critical skills.

5. The learning disabilities identification system in practice in most U.S. schools does not allow the identification of students with learning disabilities until measures of academic achievement can demonstrate a severe discrepancy between measured potential and academic achievement. Typically, this is not possible until second or third grade, past the time when intensive instruction has the best chance of remediating students' early deficits. In short, the system is problematic because we have delayed intervention to a point where it may be less effective and until we are playing catch-up.

As of 2004, the federal government is actively looking at making changes to the definition of learning disabilities in the federal law. A national center for research on learning disabilities was established in 2002 that is a collaborative effort between Vanderbilt University and Kansas University. One task that this center has undertaken is the examination of alternative models for the identification of learning disabilities. An emerging approach to identification of learning disabilities is called response to intervention, or RTI. In short, RTI is an approach in which individuals with learning problems often experienced by students with learning disabilities are placed in high-quality treatment, research-validated intervention programs across relatively short periods of time. The student's response to these interventions is monitored. Students whose performance is both far below age/grade-level peers and who do not respond positively and sufficiently to the validated treatment approaches may be considered for learning disabilities services. Much research is ongoing with regard to RTI approaches to identification of learning disabilities, and it is anticipated that successful implementations of RTI will inform and influence the learning disabilities definition that will be a part of federal law in the very near future.

3. TREATMENTS FOR LEARNING DISABILITIES

As important as defining learning disabilities has been the process of conducting research on effective treatments. The literature on treatment for learning disabilities is broad and extensive, so this summary will necessarily be abbreviated. To focus this section further, primarily school-based treatments will be discussed, though individuals with learning disabilities experience challenges beyond the school setting. Finally, instead of attempting to chronicle every researched school-based approach, this summary is organized around three topics: treatments without widespread research support, general principles undergirding generally effective treatments, and selected examples of effective approaches for educating students with learning disabilities. At a general level, the treatments reported in this section reflect well-established intervention methodologies with significant research examining their effectiveness.

3.1. Treatments That Do Not Have Widespread Research Support

3.1.1. Modality Instruction

Modality instruction is the practice of assessing certain characteristics of the learner (i.e., modalities), such as whether the individual is a visual, auditory, or kinesthetic learner, and then matching treatments to the individual's strengths. It has long been thought that matching treatment to measured processing strengths can provide one mechanism to improve learning for persons with learning disabilities. The assumption is that if we teach individuals through their preferred modality, then they will learn more and/or better than they would if we did not tailor our instruction in this way. These approaches, sometimes called aptitude-by-treatment interactions or diagnostic-prescriptive teaching, have great intuitive appeal. They sound logical and reasonable. Unfortunately, the effectiveness of these approaches has not been empirically proven on a wide scale. For these methods to work, we need to be able to demonstrate that individuals with certain characteristics learn better with instruction matched to their strengths. These effects have been very difficult to find. Instead, treatments that are generally effective for students with learning disabilities also tend to be effective for students without learning disabilities.

3.1.2. Psycholinguistic Training

Psycholinguistic training is an approach to training people in processes that they are believed to be deficit in. It was at one time a major intervention approach for students with learning disabilities. Training approaches were often based on results from a very popular test called the Illinois Test of Psycholinguistic Abilities (ITPA). This test measured integrative, receptive, and expressive linguistic abilities through presenting test subjects with information through visual and auditory channels. The assumption was made in this treatment approach, as in others, that discrete psycholinguistic abilities can be measured directly and then remediated. This assumption has been the subject of intense research scrutiny throughout the 1970s and 1980s. There were those who claimed with data that psycholinguistic training was generally not effective, and others with data who refuted these claims, stating that psycholinguistic training provided discrete benefits. The claims and counterclaims become a bit confusing. However, based on a review of the existing literature by this author, it appears that at least some of the areas measured by psycholinguistic assessments can be enhanced by psycholinguistic training. Particularly in the "expressive" areas of manual expression and verbal expression, evidence appears to confirm that these areas can be improved moderately through direct training. In the other 10 subareas measured by the ITPA, it appears that the effects of training are more modest.

Despite these moderately positive findings, a number of researchers have questioned the practical utility of gains of the magnitude reported. That is, it may be possible that gains in specific psycholinguistic variables are possible. The question is, do these gains translate into important gains in functioning in other areas of a person's life such as reading or language use? In the absence of these types of evidence, we must question whether these psycholinguistic interventions should be the highest priority for persons with learning disabilities.

3.1.3. Perceptual Motor Training

Probably the most frequent form of "process training" in special education has been the use of what is generally called perceptual motor training. The theory goes that children with learning disabilities, or at least some of them, have problems in sensory integration to which some or all of their learning problems can be attributed. These problems may include disorders in the postural, ocular, and bilateral integration, a disorder in the ability

to plan and execute motor acts, a disorder in form and space perception, auditory language problems, and/or tactile defensiveness. Numerous clinical measures intended to measure these problems have been created over the years. For children with such problems, the idea is that direct sensory motor training could remediate these children's underlying problems, and hence mitigate the effects of their disability. There have been many testimonials to the effectiveness of perceptual motor training, and a number of clinical trials have fanned the popularity of these treatments. Research support on a broad scale, however, has not emerged for these therapies. It may be that persons with learning disabilities have problems in these areas, but consistent, effective treatments that result in consistent, large-scale positive effects have not been identified and validated on students' perceptual/sensory motor skills, their academic achievement, or their cognitive performance or aptitude.

3.1.4. Placement in Special Education

Perhaps the most frequent treatment for children with learning disabilities in U.S. schools is placement in special education (as allowed by the IDEA). It is difficult to make blanket statements about special education as an enterprise, because so many different practices exist in these settings. Moreover, there are methodological problems associated with measuring the overall effectiveness of such a diverse enterprise. Despite these challenges, a number of studies have examined the effectiveness of special education placement for students with mild disabilities such as learning disabilities. In general, these studies have found weak effects for special education placement as a treatment. These findings, however, must be tempered by the awareness that many of the research-based effective practices identified in the last section of this article are most frequently carried out in special education placements. As such, it is probably most accurate to say that practices in special education are effective or ineffective, as opposed to placement. The extent to which any particular placement will be effective or ineffective for students with learning disabilities is a direct result of the effectiveness of practices used in that setting.

3.2. Characteristics of Generally Effective Instructional Interventions

Researchers have identified a series of components that characterize generally effective instructional interventions. These components do not guarantee that a specific program or approach will be effective with individuals with learning disabilities. However, when these principles are implemented in classrooms, children generally learn better and more. Some of the principles of effective instruction are listed below.

3.2.1. Teacher Organization

Instructional groups should be organized for efficiency and sustained focus on the instructional content. This means that students in small group instruction should be seated so that the teacher can work with each student, transitions from one instructional activity to another should be quick (which may require instruction on how to transition), and lessons should be started quickly at the beginning of lesson time.

3.2.2. Student Engagement

Students should not just be attending to academic material, but should be actively responding to material that is geared to their instructional skills. Students should be placed in materials that are challenging, in which they can respond briskly and in which they will not make large numbers of mistakes. They should experience high rates of success in the materials.

3.2.3. Opportunities to Respond

Students should receive many opportunities to respond during instruction and should get clear feedback on the correctness of their performance from the teacher.

3.2.4. Overlearning

Skills should be taught to a level of overlearning, with new skills being introduced gradually. That is, students need to be taught to a level at which they are able to perform skills almost automatically, without great effort.

3.3. Treatments for Learning Disabilities with General Research Support

In addition to treatments that have not demonstrated high levels of effectiveness and broad research support, there are a host of treatments for learning disabilities that have been found to be generally effective. This section reviews some of the most widely used and

familiar treatments in schools for learning disabilities. The list is not exhaustive, but is intended to highlight some of the most effective approaches.

3.3.1. Early Intervention

The goal of early intervention is the early detection and remediation of negative effects of disability and the prevention of developmental and learning problems for children who are at risk of developing these problems. There have been some very impressive results reported from the implementation of early intervention programs, though methodological issues have made it difficult to make large-scale, conclusive statements. In general, however, children who participate in early intervention programs have demonstrated gains in IQ, language development, motor development, and academic achievement.

One of the most exciting recent developments in learning disabilities is improvement in our understanding of early language development and early literacy. At the same time, assessment technologies have been developed that allow us to identify and remediate the problems of students with learning disabilities early, before they grow into larger ones. With these developments, learning disabilities advocates and professionals have moved one step closer to realizing their longstanding desire to provide meaningful, effective treatments for persons with learning disabilities early, when interventions can be maximally effective. We now know many critical "building blocks" of early reading and we have direct measures of these skills and each is directly teachable. Initial results for directly teaching children these building block skills appear promising. That is, many children who are at risk of learning failure seem to maintain and build upon their early gains after experiencing direct and focused reading instruction early in their educational careers.

3.3.2. Direct Instruction

Direct instruction is a method of teaching that was developed by Siegfried Engelmann and Wesley Becker at the University of Oregon. These researchers married the positive effects seen in using behavior modification with a systematic analysis of how to teach different knowledge forms (e.g., facts, rules, concepts). To this knowledge, they applied the principles of effective instructional design and developed a series of curricula for teaching academic skills that have widespread and powerful research supporting their effectiveness. These include the DISTAR and Corrective Reading series, among others. These programs provide explicit, step-by-step guidance for teachers, strategies for correcting student errors, gradual removal of teacher direction as students become more independent, the use of adequate, systematic practice with many different examples, and consistent review of recently learned materials. They have also generated some of the most widespread and consistent research support on program effectiveness of any instructional methodology.

3.3.3. Reading Comprehension Strategy Instruction

Because many students with learning difficulties have problems with language variables that often translate into reading difficulties, a large number of studies have examined the effectiveness of teaching reading comprehension strategies to students with learning disabilities. Many different approaches have been examined, including the use of advance organizers, vocabulary interventions, the use rules to help comprehend text, self-questioning, the use of previewing and reviewing strategies, and memory enhancement strategies, to name a few. Though the effectiveness of different comprehension strategies varies, the results are generally in the good to excellent range. These findings are especially encouraging because they collectively demonstrate that reading comprehension can be enhanced significantly for students with learning disabilities and that a wide variety of strategies are available that can have significantly positive effects.

3.3.4. Behavior Modification

Behavior modification is an approach to treatment that is based on the science of applied behavior analysis. Basically, this approach works through careful application of behavior change principles to the problems experienced by students with learning disabilities. In behavior modification, the behaviors that are the subject of change are defined precisely, very explicit and complete instructional procedures are used to teach new behaviors, the results are measured directly, and teaching procedures are modified to produce optimal effects. People who use behavior modification consider academic performance to fall in the realm of behavior just as they view social behavior. Indeed, behavior modification procedures have a very strong research base, underlying their use for both helping improve social behavior and developing academic

skill. Behavior modification has demonstrated consistent and strong relationships to improved student performance.

3.3.5. Peer Tutoring

Peer tutoring is a set of strategies in which students serve as both teachers and learners. In general, students work with each other in a 1-on-1 setting, learning together. Typically, students serving as "teacher" at any given time follow directions regarding what and how to interact with the student serving as the "learner." Peer tutoring has the advantage that it can individualize instruction for students based on individual needs. In many cases, all students participating in peer tutoring get to serve both as instructors and as learners. A range of positive effects have been documented in the peer-tutoring literature, including positive effects on academic achievement in reading, mathematics, and language, increases in students' attitudes toward school, and improved school behavior. Gains have been documented for both tutors and tutees.

3.3.6. Progress Monitoring, Graphing, and Data-Based Decision Making

One of the few universals that can be identified at the current state of our knowledge about individuals with learning disabilities is that we cannot predict with certainty which strategies will work prior to trying them. As a result of this situation, much research and development has been devoted to identifying ways to measure growth and development for individuals with learning disabilities. The process of monitoring progress across time, graphing data, and using these data for the purpose of determining effectiveness of treatments is called formative evaluation. Recall that development of formative evaluation tools was a major purpose for one of the federally funded institutes for research on learning disabilities in the late 1970s and early 1980s. Not only have progress-monitoring and decision-making tools been developed that are specifically useful with students with learning disabilities (e.g., curriculum-based measurement), but also research has proven that using these technologies to give teachers feedback on instruction improves student performance significantly. The idea is that if teachers can get feedback on the effectiveness of their instruction over a short period of time (say 6 weeks), they will be in a good position to either keep effective instructional strategies or move on to other potentially effective strategies if what they are doing is not working. Progress monitoring, graphing, and data-based decision making provide teachers these tools.

3.4. Maximizing Results for Persons with Learning Disabilities

In general, the specific problems associated with learning disabilities are lifelong, though many of the problems can be attenuated by instruction and accommodation. Students with learning disabilities can learn to read and can become functional readers. Students with learning disabilities who have memory problems can learn specific strategies to assist with remembering. Students with learning disabilities with organizational problems can learn strategies to assist with planning and keeping themselves organized. To maximize their functioning, however, individuals with learning disabilities may need a variety of supports throughout life.

During the school years, the problems associated with learning disabilities impact students' performance in academic skill development, primarily in reading, language arts, and mathematics. In turn, students' grades and academic achievement, two important success indicators in schools, are affected. As students mature and transition into adult life, the problems they experience can affect their lives in different and important ways. Low levels of academic achievement can diminish students' likelihood of pursuing college, which can translate into less lucrative employment , less independence, or even lower standards of living. Specific skill deficits can preclude access to or participation in certain professions, and for some, deficits in social skills can make it difficult to maintain steady employment. For many individuals with learning disabilities, however, some of the most negative effects associated with their disabilities can be mitigated through instruction, advocacy, and advanced planning. Indeed, there are many, many persons with learning disabilities who have minimized the effects of their disabilities and have excelled personally and professionally.

To maximize the chances of success for students with learning disabilities, a number of scholars have argued that it is critical that increased attention be paid to the issue of preparing students for adult life while they are in school. Supports need to include vocational instruction and experiences, assistance in finding and securing employment, assistance in finding and participating in post-secondary education options, and assistance in learning to use community resources to the students'

advantage. For optimum effectiveness, transition services should include students themselves in planning their transition from school to adult life. Specifically, these services should include the student's family, should begin early (some advocate as early as elementary school), should be sensitive to familial and cultural factors, and must be comprehensive. It is through this careful planning and service delivery that successful results for individuals with learning disabilities can be achieved.

See Also the Following Articles

Behavioral Assessment in Schools ■ Behavioral Observation in Schools ■ Emotional and Behavioral Problems, Students with ■ Learning ■ Mathematics, Teaching of ■ Neuropsychological Assessment in Schools ■ Reading, Teaching of ■ School Discipline and Behavior Management ■ Teaching Effectiveness ■ Writing, Teaching of

Further Reading

Bradley, R., Danielson, L., *et al*. (Eds.) (2002). *Identification of learning disabilities: Research to practice*. Mahwah, NJ: Lawrence Erlbaum Associates.

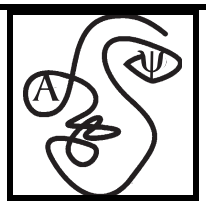
Hamill, D. D. (1990). On defining learning disabilities: An emerging consensus. *Journal of Learning Disabilities*, *23*, 74–83.

Kirk, S. A. (1975). Learning disabilities: Selected ACLD papers. In S. A. Kirk, and J. M. McCarthy (Ed.), *Behavioral diagnosis and remediation of learning disabilities*. Boston, MA: Houghton Mifflin.

Lyon, G. R. (1994). *Frames of reference for the assessment of learning disabilities*. Baltimore, MD: Paul H. Brooks.

Mercer, C. D. (1997). *Students with learning disabilities* (5th Ed.). Upper Saddle River, NJ: Merrill.

Ysseldyke, J. E., Algozzine, B., *et al*. (1982). Similarities and differences between low achievers and students labeled learning disabled. *Journal of Special Education*, *16*, 73–85.



Learning Styles and Approaches to Studying

Noel Entwistle
University of Edinburgh, Edinburgh, United Kingdom

Elizabeth Peterson
University of Auckland, Auckland, New Zealand

1. Historical and Conceptual Introduction
2. Learning Styles
3. Learning Strategies and Approaches to Studying
4. Measurement Instruments
5. Relationships to Achievement and to Learning Environments
6. Other Applications in Education and Business
Further Reading

Learning styles are relatively consistent preferences for adopting particular learning processes, irrespective of the task or problem presented but generally relating to educational or workplace settings. Learning strategies and approaches to studying are specific to education and training, being more context-specific ways of tackling learning tasks that involve characteristic combinations of intentions, learning processes, and study activities.

GLOSSARY

approach to studying A context- and content-specific way of carrying out academic tasks.
cognitive style A habitual and distinctive way of attending to, and processing, perceptual and cognitive information.
disposition A composite concept bringing together ability, inclination, and sensitivity to context.
learning strategies The way of learning adopted in tackling more specific types of tasks.
learning style A preferred and relatively consistent way of learning, usually related to educational or workplace settings.
study orientation A habitual way of tackling tasks presented in educational contexts.
teaching–learning environment The instructional activities, learning resources, assignments, and assessments set up to encourage and support learning.
thinking style A preferred way of thinking, described in general terms to apply to various contrasting situations.

1. HISTORICAL AND CONCEPTUAL INTRODUCTION

The term "cognitive style" was first used in the psychological literature during the 1940s to describe individual differences in perceptual ability, especially Witkin's identification of "field independence," that is, the ability to separate an object visually from its background. The meaning of cognitive style has since been widened to refer to other strong, relatively consistent individual differences in the way in which people cognitively process, as well as perceive, information.

Conceptually, cognitive styles have links to both personality and ability. Like personality traits, they are relatively fixed characteristics that affect the way in which people respond across a variety of situations, but their theoretical and empirical underpinnings are much less secure. Like abilities, cognitive styles affect

©2004 Elsevier Inc.
All rights reserved.

mental processing, but instead of indicating the content, speed, and effectiveness of mental operations, they describe characteristic processes through which information is handled. In addition, the level of ability (measured on a low-to-high scale) will affect performance, whereas cognitive style is indicated by the position on a bipolar scale, with each pole being equally adaptive but for rather different kinds of learning.

The term "thinking style" represents a concept intermediate between cognitive style and learning style but is seen as having broader applicability than the latter. Distinctive differences between convergent and divergent thinking scores were recognized as contrasting modes of thinking, equivalent to styles, but the term "thinking style" has more recently become associated with the work of Sternberg through his theory of mental self-government.

During the 1960s, the term "learning style" was introduced to describe individual differences in the more specific preferences shown by people learning in educational and workplace settings. Work focusing on study methods and study strategies began earlier within educational psychology, but the term "approaches to learning" emerged from educational research during the mid-1970s and its meaning was subsequently broadened to cover approaches to studying.

Generally speaking, "styles" describe relatively stable personal preferences in how cognitive processing or learning is undertaken, whereas "strategies" and "approaches" are contextually more specific, being influenced by the content and setting within which learning and studying take place.

2. LEARNING STYLES

The early work on cognitive style, led predominantly by Witkin, focused on individual differences in perceptual processing, but Witkin's later work also distinguished broader "articulated" and "global" ways of thinking that influenced more recent work on styles. A vast array of terms has been accumulated in the attempt to describe distinctive preferences, with those associated with education and the workplace being known as learning styles. By 1976, Messick had already found 19 different sets of categories, and more styles have been introduced since then. This vast array of competing conceptualizations, with no widely agreed theoretical framework to connect them, makes it difficult to single out specific contributions, but the work of Kolb was influential in the early conceptualization of learning style. Kolb identified two dimensions: one differentiating between preferences for either convergent or divergent thinking and the other indicating the relative reliance on the Piagetian mental processes of assimilation or accommodation. Relative closeness to each of the poles of the two dimensions was used to suggest four stylistic "types" that were labeled accordingly.

The subsequent proliferation of concepts describing learning styles, each with its own related learning style instrument, has led to confusion in the essential meaning of the term. However, it is possible to make sense of the literature from integrative reviews in terms of a few recurring dimensions, although with different labels. Recent reviews by Sternberg and Riding, among others, suggest that the main learning styles cover either individual preferences for a particular modality of sensory input (e.g., visual, auditory, kinesthetic) or different forms of mental representation in learning. These latter styles are exhibited in terms of preferences for one or the other pole of dimensions described as analytic/wholist, concrete/abstract, or verbal processing/visual imagery. There has been some suggestion that aspects of learning style may reflect differences in hemispheric brain functions, but the evidence for this is still contested.

Another way in which learning style constructs can be systematized usefully is by contrasting those that focus on the learning process itself (e.g., Pask's version of analytic/wholist, which he labeled serialist/holist) with others that look at instructional and environmental preferences (e.g., the Dunns' extensive array of categories related to classroom contexts).

3. LEARNING STRATEGIES AND APPROACHES TO STUDYING

Two distinct research traditions have developed for defining and measuring learning and study strategies: one applying preexisting psychological constructs and the other developing concepts from studies within the specific educational context. However, the most recent work can be seen as merging those traditions. After early attempts at using Likert scales to measure study methods and motivation, research in the United States drew heavily on cognitive psychology to describe students' typical ways of tackling academic tasks, being particularly influenced by the distinction between active elaboration of information received and routine memorization of it. During the early 1980s, the work of Weinstein brought together these ideas with

practical advice on effective studying to link the measurement of study strategies with courses designed to help develop students' studying. Around the same time, Schmeck also used insights from cognitive psychology to develop an inventory of learning processes that described deep and elaborative processing as well as fact retention and methodical study.

More recently, Pintrich introduced several additional motivational constructs into the measurement of study strategies, including aspects of expectancy, value, and affect. Expectancy refers to students' beliefs about whether they can perform a task and involves self-efficacy and control beliefs about learning. The value component indicates why students engage in particular academic tasks and is covered partly by contrasting intrinsic and extrinsic goal orientation and partly by task value. This latter concept introduces affect by exploring the extent to which students find a particular task interesting, useful, and important. Learning strategies, as in the earlier work, focus on elaboration and are seen in relation to explicit and detailed metacognitive activities and self-regulation as well as volition. Interesting interplay among expectancies, motivation, and self-regulation, with consequent influences on achievement, has been reported in the most recent work, with tentative suggestions about possible gender and ethnic variations.

Whereas much of the research on learning strategies in the United States has increasingly been extended to work in schools, the alternative tradition, describing approaches, developed in Europe and Australasia and focused mainly on university students. The early work of Biggs was influenced by information processing theory, but most of the subsequent work has established conceptual frameworks out of interview and inventory studies carried out specifically in the higher education context. Using a distinctive form of qualitative analysis of interview data, Marton and his research team found a marked contrast among students in how they described reading an academic article in a naturalistic experiment. They distinguished a deep approach (which involved reading with the intention of understanding) from a surface approach (which relied on routine memorization to "get by"). The researchers argued that these approaches were specific to a particular academic context and were also dependent on the content of the learning task. However, the successful introduction of inventory measures indicated that approaches had a certain consistency, at least across similar learning environments and subject matter. Some researchers have suggested that such typical approaches should be referred to as "study orientations," but there is still terminological disagreement in the literature.

Subsequent work by Entwistle and Biggs, working independently, described approaches to learning in terms of a combination of intention (or motive) and processes. For example, the deep approach involved the motive to understand for oneself combined with learning processes that involved an alternation between holist and serialist thinking. These descriptions were based partly on interviews and partly on factor structures of specially designed inventories, but it became clear that a third approach was necessary to explain the variations in studying reported. The strength of students' concern about their grades leads to the adoption of a strategic or achieving "approach to studying." The strategic approach seems to draw on either achievement motivation or a sense of responsibility for achievement, while the processes involve organized studying, effort, and attempts to monitor the effectiveness of studying.

This work was developed further by Vermunt, who introduced additional components relating to students' beliefs about the nature of knowledge, aims, and goals as well as to the balance between external regulation and self-regulation in studying (or the absence of regulation). This trend—seeing learning strategies in relation to epistemological beliefs—can also be found in the work of Pintrich and others in the United States. However, the most recent work by both Vermunt and Entwistle and their colleagues is distinct from the research on learning strategies in the United States in creating constructs that are composites. These cover a series of concepts that hang together so consistently that separating the component parts tends to destroy the sense of the experience reported by students. In this sense, they act rather like the "dispositions," described by Perkins, that bring together ability, inclination, and sensitivity to context.

4. MEASUREMENT INSTRUMENTS

A wide and disparate range of learning style inventories exists. This is due in part to the failure to reach any strong consensus on the theoretical basis of learning styles. Thus, the instrument that is chosen will depend largely on the context to which it is applied and the type of learning style that the investigator wants to assess. Furthermore, as Coffield and colleagues recently argued, the inventories vary considerably in how much evidence of reliability and validity is provided and in the extent to which the evidence is independent of the originator of

the inventory. This is particularly problematic where the inventories have been vigorously marketed, with strong claims for their value in educational or workplace contexts yet without convincing psychometric substantiation. In contrast, there is greater agreement and fewer inventories describing the main learning strategies and approaches to studying, and the psychometric properties of the most fully developed inventories have been firmly established.

However, uncertainties about the instruments used in this area of research do remain because the majority of instruments have been operationalized in self-report formats, usually Likert scales. Thus, the validity of these instruments depends on the extent to which individuals are conscious of the strategies they typically use in their learning and studying and can accurately reflect on them. The items are sometimes quite transparent in meaning and value loading, and students can, to some extent, present whatever impressions they like. However, the style items tend to offer alternatives that can be viewed as equally attractive, and the factor structures of the strongest learning strategy and approaches to studying inventories are sufficiently consistent across both subject area and educational system to substantiate their validity claims. Some instruments have adopted forced-comparison formats or ratio scoring, but both of these methods create difficulties for conventional statistical analysis. Messick recently argued that the uncertainties of measurement that exist in the measurement of styles necessitate the use of normative and ipsative scoring techniques in combination as well as the collection of data from contrasting populations.

5. RELATIONSHIPS TO ACHIEVEMENT AND TO LEARNING ENVIRONMENTS

The underlying purpose in using both learning style and approaches to studying concepts in educational research is to find ways of helping learners to become more aware of their styles and approaches and, thus, to seek ways of improving the effectiveness of their learning and studying. Evidence for the validity of the inventories rests, in part, on the relationships that exist between the scale scores and the measures of academic achievement. The unconvincing or inconsistent relationships often obtained are not surprising given the psychometric problems with several of the style inventories, but these results could just as well be attributable to the varied definitions of achievement and its sometimes unreliable measurement in educational institutions. However, subscales of both learning strategy and approaches to studying inventories do relate to student achievement more convincingly but are still dependent on the type of learning being assessed. Although few studies have reported consistent gender differences, suggestive cultural and ethnic variations are emerging.

Theoretically more interesting research has been relating both style and approaches or strategies to teaching and learning in various subject areas, contrasting learning environments, and different educational stages. For example, in the style literature, there has been considerable debate about whether it is better to match or mismatch learner preferences with teaching preferences, how this should be done, who benefits from this, and who should decide from among the alternatives. Few large-scale, independent studies have been done to investigate this, but the findings do suggest that when the styles of the teacher and the learner are matched, learning is perceived as more congenial and teachers are rated more positively.

In the approaches literature, it is becoming clear that a deep, self-regulated approach can be encouraged through methods of teaching and assessment that focus explicitly on conceptual development and subject-specific thinking skills and that also require students to reflect on their own ways of studying. But the most recent work from both Biggs and Entwistle has stressed the necessity of thinking about the influence of teaching–learning environments on learning in terms of interacting systems. The students' approaches to studying and perceptions of teaching are part of one system in which the separate constructs used by researchers artificially separate aspects that are integrated within the individual and have their origins in earlier experiences. The idea of a "disposition to learn through personal understanding" would bring together ability, conceptions, intentions, style, motivation, strategies, and approaches, seen as constructs with different levels of generality and relative stability that work together to influence an individual's study behavior. The teaching–learning environment is a system that can be described at different levels as well, concentrating on the specific content and teaching methods devised by teachers, the broader social experiences of students, or the institutional and sociopolitical context within which education takes place. All of these

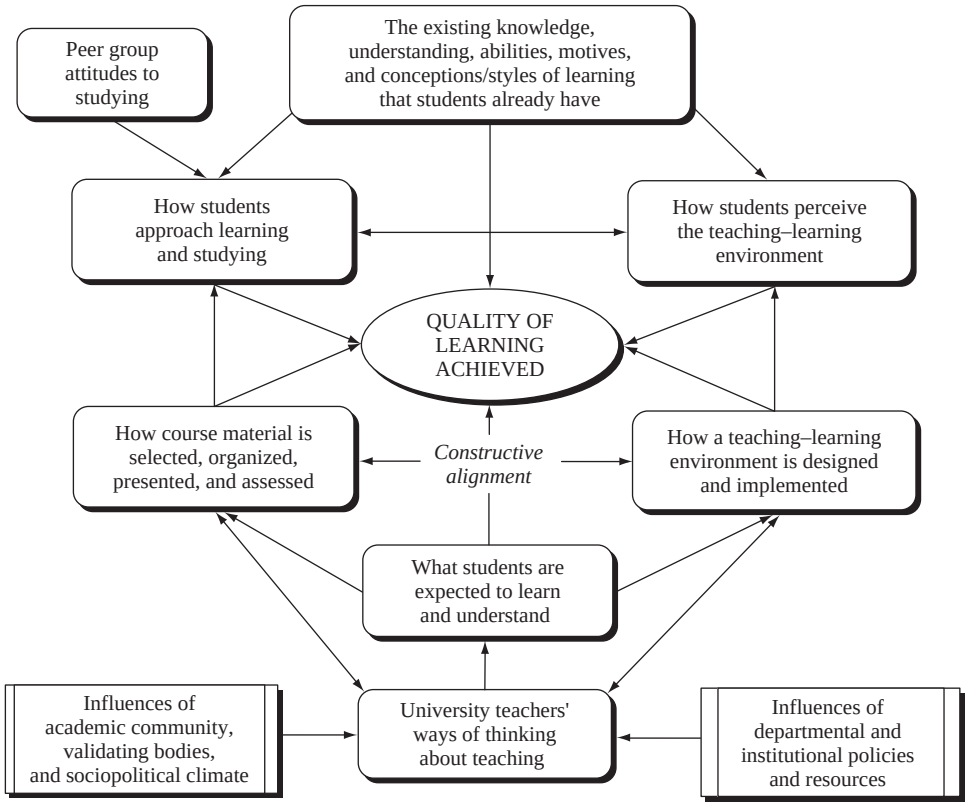

Peer group attitudes to studying
The existing knowledge, understanding, abilities, motives, and conceptions/styles of learning that students already have
How students approach learning and studying
How students perceive the teaching–learning environment
QUALITY OF LEARNING ACHIEVED
How course material is selected, organized, presented, and assessed
Constructive alignment
How a teaching–learning environment is designed and implemented
What students are expected to learn and understand
Influences of academic community, validating bodies, and sociopolitical climate
University teachers' ways of thinking about teaching
Influences of departmental and institutional policies and resources

FIGURE 1 Interacting systems affecting the quality of student learning.

act together to influence student learning, but their influence is not uniform due to the different ways in which individual students react to the environment they experience. Figure 1 suggests what such interacting systems might look like in an outline based on ongoing research into teaching–learning environments in higher education.

However, the interacting systems do not show which aspects of each component of the environment are most likely to promote "deep learning." For that, one must bring together a series of studies in schools such as Project Zero at Harvard University on "teaching for understanding" and de Corte and colleagues' study in Belgium on "powerful learning environments." Work at the university level also contributes to the listing shown in Table I, drawing on recent studies by Biggs in Australia on "constructive alignment" and by Entwistle, Hounsell, and their research team in the United Kingdom on ways of "enhancing teaching–learning environments."

TABLE I
Teaching Activities Encouraging a Deep Approach to Learning

- Specifying broad generic aims emphasizing understanding and using "throughlines" to remind students of those aims on a regular basis
- Ensuring that teaching methods and assessment procedures work in consort to support those aims, thereby providing constructive alignment
- Using instructional methods that provide clear conceptual explanations, are presented with enthusiasm, and demonstrate empathy with the learners
- Providing authentic tasks and realistic problems that have personal meaning and involve understanding performances
- Using group discussions of both the content and the process of learning and studying
- Promoting students' awareness of their own cognitive processes and approaches to studying as well as their ability to control their motives, feelings, and efforts

6. OTHER APPLICATIONS IN EDUCATION AND BUSINESS

In spite of the lack of convincing evidence about the predictive validity of learning style inventories, they have been used widely in both education and the business world because the underlying conceptualizations tend to make intuitive sense to the users. The process of reflecting on the meaning of the scores then encourages metacognitive activity about learning preferences. However, learning style inventories have also been used for selection and promotion in companies, where any lack of convincing evidence about the instruments' psychometric properties and predictive validity has serious consequences.

Learning strategy and approaches to studying inventories have been used to identify students who reveal ineffective study strategies as well as to monitor the effects of innovations in teaching and learning. Although these inventories do seem to provide valid findings at the group level, their use in guiding individuals is much more problematic and needs to be supplemented by individual discussion of the responses prior to any changes in approach being suggested. Experience in using the research findings to encourage more effective studying strongly suggests that stand-alone teaching of study techniques is much less effective than building repeated suggestions and advice into the curriculum itself, thereby making them specific to the tasks that students undertake on a regular basis.

See Also the Following Articles

Achievement Tests ■ Assessment and Evaluation, Overview ■ Cognition and Culture ■ Cognitive Skills: Training, Maintenance, and Daily Usage ■ Educational Achievement and Culture ■ Educational and Child Assessment ■ Intelligence Assessment ■ Learning ■ Mathematics, Teaching of ■ Psychometric Tests ■ Reading, Teaching of ■ Teaching Effectiveness ■ Writing, Teaching of

Further Reading

Biggs, J. B. (2003). *Teaching for quality learning at university*. Buckingham, UK: Open University Press.

Coffield, F., Ecclestone, K., Hall, E., Meagher, N., & Mosely, D. (2004). *A systematic and critical review of the literature on learning styles and pedagogy in post-16 learning*. London, UK: Learning and Skills Development Agency.

Collis, J. M., & Messick, S. (2001). *Intelligence and personality: Bridging the gap in theory and measurement*. Mahwah, NJ: Lawrence Erlbaum.

De Corte, E., Verschaffel, L., Entwistle, N. J., & van Merrienboer, J. (Eds.). (2003). *Powerful learning environments: Unravelling basic components and dimensions*. Oxford, UK: Elsevier Science.

Dunn, R., & Griggs, S. (2003). *Synthesis of the Dunn and Dunn learning styles research*. New York: St. John's University Press.

Entwistle, N. J. (2000). Approaches to studying and levels of understanding: The influences of teaching and assessment. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research* (Vol. 15, pp. 156–218). New York: Agathon Press.

Entwistle, N. J. (in press). *University teaching and student learning*. London: Palgrave Macmillan. (See also the project Web site at www.ed.ac.uk/etl)

Jonassen, D. H., & Graboski, B. L. (1993). *Handbook of individual differences, learning, and instruction*. Hillsdale, NJ: Lawrence Erlbaum.

Kolb, D. A. (1984). *Experiential learning: Experience as a source of learning and development*. Englewood Cliffs, NJ: Prentice Hall.

Messick, S. (1996). Bridging cognition and personality in education: The role of style in performance and development. *European Journal of Personality*, *10*, 353–376.

Pintrich, P. R., Brown, D. R., & Weinstein, C-E. (Eds.). (1994). *Student motivation, cognition, and learning*. Hillsdale, NJ: Lawrence Erlbaum.

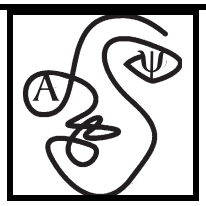
Riding, R., & Rayner, S. (1998). *Cognitive styles and learning strategies*. London: David Fulton.

Schmeck, R. R. (1988). *Learning strategies and learning styles*. New York: Plenum.

Sternberg, R. J., & Zhang, L-F. (Eds.). (2001). *Perspectives on thinking, learning, and cognitive styles*. Mahwah, NJ: Lawrence Erlbaum.

Vermunt, J. D. (1996). Metacognitive, cognitive, and affective aspects of learning styles and strategies: A phenomenographic analysis. *Higher Education*, *31*, 25–50.

Wiske, M. S. (Ed.). (1998). *Teaching for understanding: Linking research with practice* [Project Zero]. San Francisco, CA: Jossey–Bass.



Legal Competency

Norman G. Poythress
University of South Florida, Tampa, Florida, USA

1. Introduction: The Structure of Mental Health Laws
2. Components of Legal Competencies
3. Contexts for the Assessment of Legal Competency
4. Issues in Forensic Assessment
Further Reading

GLOSSARY

adjudicative competence Functional legal abilities that a criminal defendant must have in sufficient degree in order for the adjudicatory process to go forward. Also called competence to stand trial, competence to proceed, and fitness to proceed.

forensic assessment instrument (FAI) An assessment test, tool, or device designed to measure or to inform a clinical judgment about specific functional legal abilities required in a given legal context.

functional legal ability A legally prescribed capacity that an individual must have in sufficient degree in order to participate in a legal proceeding (e.g., to face prosecution in a criminal court) or to avoid legal intervention into his or her affairs (e.g., to manage one's own financial affairs or to make decisions about one's medical treatment independent of a court-appointed guardian).

mental state at the time of an offense (MSO) General term inclusive of various legally recognized mental states (e.g., insanity, diminished capacity, and automatism) that provide the bases for a defense to culpability for criminal behavior.

Legal competence refers to a legally prescribed ability, or set of abilities, that a person must have in sufficient degree in order to enjoy, without government intervention, the rights, benefits, and opportunities to which citizens are ordinarily entitled. Issues of legal competence generally arise in discrete legal contexts (e.g., during police interrogation or court procedures), although some may arise in clinical (e.g., competence to make decisions about one's own medical treatment) or research (e.g., competence to consent to research participation) contexts. The conceptual and contextual framework for assessments of legal competence differ from those that guide more traditional therapeutically oriented assessments, and an array of forensic assessment instruments have been developed to assist mental health professionals whose evaluations address these legally focused competencies.

1. INTRODUCTION: THE STRUCTURE OF MENTAL HEALTH LAWS

In the United States and many other democratic societies, the government cannot arbitrarily deny citizens the rights, benefits, and opportunities to which they are entitled. However, under either of two legal doctrines, citizens may be deprived of such rights and benefits if it is determined that they lack the requisite mental capacity to exercise such rights. One legal doctrine is police protection, which permits the

©2004 Elsevier Inc.
All rights reserved.

government to interfere in the lives of citizens whose behavior poses too substantial a risk to their own safety or the safety of others. This doctrine underlies, for example, involuntary hospitalization laws that permit the government to impose mandatory psychiatric treatment on individuals who pose a substantial risk of harm to others. The other doctrine is the parens patria doctrine, in which the motive for state intervention is a paternalistic one. Procedures such as the appointment of legal guardians are intended to protect individuals from adverse outcomes in day-to-day situations that might result from their own poor judgments that stem from mental impairment. This doctrine underlies, for example, procedures that provide for the appointment of legal guardians to manage a person's personal and/or financial affairs.

Intrusions into citizens' lives on the basis of these rationales are prescribed in a collection of laws, broadly conceived as mental health laws. Associated with each of these laws is a legal competency that people must have in order to avoid state intrusion and to retain the full array of rights to which citizens are ordinarily entitled. As Morse elucidated most clearly in 1978, the structure of such laws involves three components: (i) a predicate mental condition that (ii) "causes" (iii) impairment in functional legal ability.

Most commonly, the predicate mental condition is described in legal terms, such as "mental disease" or "mental defect." Although in most instances there is no direct translation of these terms into contemporary psychological or psychiatric nosology, for more than a century mental health professionals have been utilized by the courts to provide evidence about the legal competence of parties to litigation. Historically, clinicians approached issues of legal competence using diagnostic interviews and general psychological tests (e.g., intelligence tests, general personality inventories, and projective tests) that they routinely employed in, and that were developed for, therapeutic applications. Today, however, the limitations of such methods for assessing legal competencies are widely recognized. Although the information gleaned from diagnostic interviews and traditional tests may be informative to judgments about the first element—predicate mental condition—these methods typically have not been normed or validated on relevant legal populations in ways that help inform the causality and functional impairment elements. In other words, the components of legal competencies involve considerations that go beyond what can be gleaned from conventional clinical assessment approaches.

2. COMPONENTS OF LEGAL COMPETENCIES

Grisso identified five features of legal competencies that distinguish them from more general competence constructs: a functional component, a causal component, an interactive component, a judgmental component, and a dispositional component.

2.1. Functional Component

The primary concern in forensic evaluations is the person's functional legal ability—the third element in the structural equation of mental health laws. As far as the law is concerned, it is the degree of impaired legal functioning, not the mere presence of mental disorder or generally impaired mental functioning, that determines whether a person's usual rights or privileges may be denied.

Several additional observations are important here. First, the particular functional abilities of concern vary from one legal context to another; to compare a couple of contexts alluded to previously, involuntary hospitalization proceedings focus on (among other things) a person's capacity to provide adequately for his or her basic needs (i.e., food and shelter), whereas guardianship proceedings focus on the capacity to manage one's personal affairs (e.g., to manage money and to enter contracts). Second, although a person's general mental condition is a relevant consideration insofar as a legally cognized predicate condition must be established, diagnosis is of secondary importance to functional legal ability. A person may have no mental disorder or several of them, but state intervention is justified not on the mere finding of mental impairment. Third, for any given legal context the relevant functional abilities are relatively narrow, and it is for this reason that some traditional clinical measures may be limited as a basis for inference about legal incompetence. Intelligence tests can provide information about a person's general problem-solving ability, but the legal questions will often be more narrowly focused and therefore require a more nuanced analysis as a basis for inference about impaired legal functioning.

2.2. Causal Component

Judgments about legal competence require causal inferences about the reasons for a person's functional abilities or deficits. Legal criteria for competence, often found in statutes governing a particular type of

proceeding or conflict (although criteria may also appear in appellate cases, administrative regulations, or other legal sources), often specify that the third element—impaired legal functioning—must be "as a result of" the predicate mental condition (first element).

Research that relates specific mental disorders, symptom patterns, neurological impairments, cognitive capacities, and so forth to functional legal abilities provides a basis for bridging the assessment of mental condition and legal competence. A further consideration in legal competence assessments is to distinguish legally relevant "causes" from irrelevant ones. For example, the legal context of adjudicative competence—more familiarly called competence to stand trial—requires (among other things) that a criminal defendant be able to adequately assist his or her attorney. For a defendant who is not able to do so because of impairments attributable to mental disease or defect, the adjudicatory process must be postponed until he or she regains such ability. However for a defendant, even one who has a mental disorder, who is merely unwilling to work with his or her attorney, the process can go forward. Assessments of legal competence may call for mental health professionals to aid the legal system in making such distinctions.

2.3. Interactive Component

Judgments regarding a person's legal competence are not dependent merely on his or her mental capacity or the absolute level of his or her functional ability. His or her particular circumstances within the legal context may also be of relevance. It is the fit or congruence between capacity and demand that is ultimately of importance. From this perspective, one can see that courts (or other legal decision makers) might reach different conclusions about the legal competence of two individuals who have the same (measured) abilities.

Consider the adjudicative competence context, for example, for two individuals of modest mental capacity (e.g., judgment, memory, and general knowledge). One might be facing very simple charges such as driving without a valid license; the only issues of contention in the case are likely to be whether he or she was behind the wheel of a car and whether he or she possessed a valid license at the time. The other individual may be charged with intention to commit securities fraud, for which the defense (and required degree of assistance to counsel) is likely to be much more prolonged and complex. A mentally disordered individual of modest functional capacity might reasonably be judged "competent to proceed" in the first case but perhaps not in the second.

As this example illustrates, it will be not only the individual's measured level of psycholegal functioning but also the match of capacity with anticipated demand that will be relevant to the ultimate legal judgment of competence. Thus, to the degree possible, assessments of legal competence may require that the mental health professional assess and take into account the anticipated demands of the situation on the defendant.

2.4. Judgmental Component

In considering the adequacy of the fit between a person's abilities and the demands of his or her legal situation, the law applies a variety of flexible legal tests. As will be evident in the next section, words such as "sufficient," "reasonable," or "substantial" are common qualifiers of the degree of capacity that the law demands.

There are at least three important consequences of the use of these qualifiers. First, these terms facilitate individualized judgments by legal decision makers. The competency criteria, although explicit, are not of the "hard-and-fast," "one-size-fits-all" variety. Such terms permit a reasoned application of complex rules in judging cases that may vary radically in their particulars. Second, these terms reflect that most legal judgments about competence also involve social and moral considerations. In a criminal case involving an insanity defense, for example, such as the notorious case of John Hinckley, the decision about the accused's mental state at the time of the offense is ultimately one of moral culpability. That is, within the framework of the written definition of legal insanity, was this particular individual's impairment of the type that would warrant his or her exclusion from moral culpability for criminal behavior?

Following from the first two observations, the third consequence is that mental health professionals are (or at least logically should be) limited in the scope of conclusions that they draw in legal competency evaluations. The social and moral issues alluded to here are of a type that go beyond and stand outside the scope of scientific inquiry. Psychology and psychiatry have no "capaci-meters" for distinguishing the moral gradations that are part and parcel of legal competency judgments. Thus, the challenge to forensic examiners in legal competence evaluations is to provide clinical information that informs the judgment about competency while stopping short of offering conclusory opinions as to what that ultimate judgment should be.

2.5. Dispositional Component

Legal outcomes or dispositions hinge on the judgmental question—tersely put, whether or not the individual is judged "competent" (or "insane" or other analogous terms). The preceding section suggests that mental health professionals should not venture categorical conclusions about the competency judgment due to the social and moral nature of such decisions. It follows, therefore, that they should not presume to tell a legal decision maker what the appropriate disposition should be in a given case. Rather, clinical input regarding dispositions is best framed using "if—then" formulations that do not constrain the range of possible legal judgments. Whether, how, and the type of clinical assessment and treatment or other intervention might be provided can be discussed separately for different hypothetical outcomes; for example, "If the court decides X, then .... If the court decides Y, then ...."

3. CONTEXTS FOR THE ASSESSMENT OF LEGAL COMPETENCY

Issues of legal competency arise in many situations, several of which have been alluded to previously. In this section, several of the more frequently occurring legal competency questions are discussed in-depth. The reader is warned that not all legal competency issues are recognized in all jurisdictions; for example, legal insanity is a recognized defense throughout most of the Western world, although in the United States there are approximately four states in which this defense is not available. Similarly, the formal legal criteria or "tests" of competency or capacity are not identical across jurisdictions. Thus, the presentation here is necessarily more conceptual and generic in nature.

3.1. Competencies in the Criminal Justice Context

Table I summarizes several issues of "competence" that arise in the criminal adjudicatory process.

3.1.1. Competence to Waive Miranda

Confessions are the primary evidentiary basis for many criminal convictions. In the United States, the legal basis for challenging the admissibility of defendants' confessions resides in the Fifth and Sixth Amendments to the Constitution, which protect the defendant from compulsory self-incrimination and guarantee the right to the presence of legal counsel at all relevant stages of the adjudicatory process. Thus, when the accused individual is questioned by the police he or she must be informed of the legal rights that apply at the time of questioning. These include the right to have an attorney present during

TABLE I
Issues of Legal Competence in the Criminal Adjudicatory Process

Competence issue	*Fundamental inquiry*	*Legal criteria (generically)*
Competence to waive Miranda rights	When giving a statement to the police, did the defendant competently waive his or her right against self-incrimination?	Cognitive issue: Was the defendant's waiver made knowingly and intelligently? Volitional issue: Was the defendant's waiver made voluntarily?
Adjudicative competence	Can the defendant competently participate in any of several phases of the adjudicatory process?	Does the defendant have a factual and rational understanding of the proceeding, and can he or she adequately assist counsel at the proceeding?
Mental state at the time of the offense	Retrospective inquiry: Was the defendant morally culpable for his or her criminal behavior?	Cognitive issues: Did the defendant know right from wrong? Understand the nature and quality of his or her acts? Volitional issues: Could the defendant control his or her behavior? Was his or her will overcome by an irresistible impulse?
Competence to be executed	Does the offender understand the nature and purpose of the sentence of death?	Effectively the same as for adjudicative competence but applied at the post adjudication phase of sentencing.

questioning, the right to have an attorney appointed for them if they cannot afford to hire their own attorney, and the right to refuse to answer questions altogether. They must also be advised that what they say during interrogation may be used in court to prove their guilt.

Challenges to the competence of a defendant's waiver are almost always made after-the-fact. Defense counsel hopes to have the defendant's prior statement ruled inadmissible at a pretrial "suppression hearing" on grounds that the defendant did not competently waive his or her rights. To the degree that the state's case depends on incriminating information obtained from the defendant, the outcome of such suppression hearings may be critical.

One key feature of clinical assessment of competence to waiver Miranda is that the evaluations are almost always retrospective. The attorney finds out days or weeks later that the defendant has given a statement to the police. The mental health expert who consults with the defense is asked to evaluate the nature and circumstances of that prior interrogation and to develop clinical information that informs the judgment of the defendant's competence at that time. Thus, the evaluation often hinges in part on archival information (e.g., transcripts or recordings of the defendant's statement) and the recollections of those present (e.g., interviews with the defendant and police) as bases for reconstructing the defendant's prior mental state.

Another feature is that the legal "test" for evaluating this particular legal competence, unlike most others, does not involve an explicit "mental disease" or "mental defect" type of predicate. Rather, the court is required to consider "the totality of the circumstances" that may have impinged on the defendant's competence to waive his or her rights, including a consideration of the appropriateness of the police behavior during the interrogation.

The criteria, as indicated in Table I, note the relevance cognitive factors—whether the waiver was made knowingly and intelligently. Personality features such as general intelligence and susceptibility to suggestion or influence of authority figures are often relevant in these cases. Similarly, whether one's capacity to make a rational or well-reasoned waiver was adversely affected by drug or alcohol intoxication or symptoms of mental illness is a relevant inquiry. Some standardized tests of Miranda comprehension, modeled after more general vocabulary and verbal comprehension tests, have been developed to assist in making a highly focused inquiry into the relevant cognitive capacities.

Although Table I also indicates that volitional issues—the "voluntariness" of a waiver—are relevant, the law in this area has developed in a way that limits severely the relevance of clinical input on this issue. The higher courts have generally held that there must be clear evidence of inappropriate behavior by the police (e.g., excessive threats or other coercive behavior and interrogation periods of unreasonable duration) in order to infer that the defendant's waiver was not of his or her own free will. Intrapersonal factors, including the presence of symptoms such as delusional beliefs, obsessive thoughts, or hallucinations that "urge" the defendant to confess, have generally not been accepted as a basis for concluding that a waiver was not voluntary. Because mental health professionals will typically have no particular expertise in judging when and whether the police have exceeded legally acceptable norms in their interrogations, their assessments are typically limited in relevance to issues of cognitive capacity.

3.1.2. Adjudicative Competence

Adjudicative competence refers to a defendant's capacity for meaningful participation in the adjudicatory process. Historically, the term *competence to stand trial* was often used, although it seems clear in retrospect that this term was inappropriately narrow. In particular, criminal justice data reveal that more than 90% of defendants resolve their cases by entering pleas rather than by going to trial; thus, competence to stand trial is literally a moot concern in the vast majority of cases, and "competence to enter a plea" is a much more relevant consideration. Although competence for both of these phases—participating in a trial or entering a plea—involves the "understand and assist" test shown in Table I, so too does competence for other adjudicatory activities, including sentencing. Thus, the more comprehensive term adjudicative competence better describes the breadth of this legal competency test.

In the United States, the legal basis for adjudicative competence questions reside in the Sixth and Fourteenth Amendments to the Constitution. The Sixth Amendment guarantees the right to the assistance of counsel and to confront one's accusers; these rights cannot be effectively implemented by an individual whose mental capacities are too eroded by "mental disease or defect." The Fourteenth Amendment, which deals with issues of fairness in legal proceedings, is also implicated. Although these legal foundations reflect the defendant's concerns, the state has its own independent rationale for ensuring that only competent defendants are prosecuted. That concern is with public opinion in the larger world

community about the kind of justice afforded citizens in this country.

Because of these multiple rationales for the adjudicative competence inquiry, questions about the defendant's competence can be raised by either party or by the court on its own motion. When such motions are granted, the adjudicatory process is temporarily suspended pending the result of clinical evaluations of the defendant's competence. Every state and the federal government make provisions for government-funded evaluations of the defendant's competence to proceed; both the state and the defense may also privately retain mental health consultants at their own expense if they want to obtain independent evaluations. When the clinical evaluations are completed, a hearing may be held before a judge or jury (depending on jurisdiction rules) to determine whether the defendant is competent to proceed. If the defendant is found competent, the adjudicatory process continues (e.g., a date for entering a plea, or a trial date, may be set). If the defendant is found incompetent, then he or she is typically referred for psychiatric treatment to "restore" competence. At such time as the defendant's competence is restored, he or she is returned to court and the adjudicatory process resumes.

Unlike competence to waive Miranda, adjudicative competence is a present state evaluation. The focus is on whether the defendant is now (and in the reasonably short-term future) able to understand the proceedings and assist counsel. Also, in most (if not all) jurisdictions it is explicit that only impairments resulting from "mental disease or defect" are legitimate bases for finding the defendant incompetent.

The relevant functional legal abilities for adjudicative competence are shown in Table I. Factual understanding concerns a defendant's awareness and comprehension of how the proceedings (e.g., entering a plea or a trial) are designed to work in principle. Relevant queries include knowledge of the proper roles of participants (e.g., "What is a judge, prosecutor, or jury supposed to do?"), the nature of the proceedings themselves (e.g., "What is the nature of a plea agreement?"), and so forth. Rational understanding is more specifically focused on the defendant's beliefs about his or her own case; that is, even if a defendant has a good "civics book" understanding of how the process is designed to work, is his or her view of how the process will play out in this particular case compromised by delusional beliefs (e.g., that God will intervene to assist in the defense)? Capacity to assist counsel comprises the ability to understand the defendant's role (i.e., as the one who stands accused and whose freedom is in jeopardy); to recognize and put forth information that will facilitate the defense attorney's efforts to construct a defense; and the ability to participate, with the degree of autonomy that the law requires, in decisions that the attorney is proscribed by law from making on behalf of the defendant (e.g., whether to enter a guilty plea and whether to testify).

A number of measures or tools, most developed with substantial input from mental health professionals, are available for use in forensic evaluations of adjudicative competence. They range widely in design and structure, from some that constitute little more than menus of topics or issues to raise with defendants to others that constitute semistructured interviews with ratings of (in)capacity on explicit dimensions (e.g., knowledge of roles of court participants and comprehension of available pleas) and measures that have standardized inquiries and interpretive norms for selected psycholegal abilities.

3.1.3. Mental State at the Time of the Offense

Retribution and deterrence are two of the primary legal justifications for adjudication and punishment of individuals who violate society's laws. However, dating back to antiquity, various cultures have recognized that it is morally inappropriate to hold certain classes of individuals fully responsible for criminal behavior. Mentally disordered individuals constitute one such class of individuals. Compassion, rather than vengeance, has been recognized as a more appropriate emotional reaction to individuals who were so mentally impaired that they "did not know what they were doing" or "could not help themselves." Furthermore, such individuals would not likely be deterred from future criminal behavior through the usual mechanisms of punishment. Hospitalization and treatment have been recognized as morally more appropriate societal responses to such individuals. Thus the rationale for the insanity defense and other mental state defenses.

The insanity defense is premised on the presence of legally defined mental disease or defect. Although there are no direct translations of recognized psychiatric diagnoses into such categories, empirically it is clear that this defense is usually reserved for individuals afflicted with the more debilitating disorders (primarily psychoses) or other acute symptom presentations that resulted in severe impairments in rational judgment (e.g., "knowing right from wrong" or capacity to control one's behavior) (Table I). The defense usually will not accommodate a defendant who was generally anxious, mildly depressed, disgruntled, or was merely "having a bad day."

The clinical inquiry into a defendant's sanity is a highly focused, retrospective evaluation. As with other legal competencies, the issue is not whether, at the time of the offense, the defendant merely had symptoms of a mental disorder that the trier-of-fact might judge constitutes a mental disease or defect. Mentally ill individuals may commit crimes for the same reasons that non-ill individuals do; having schizophrenia or bipolar disorder does not insulate one from the avarice that can motivate theft or robbery, nor do such disorders protect one from aggression based on anger or vengeance.

Thus, the critical linkage is the causal relationship (if any) between the symptoms of mental disorder and the alleged criminal behavior, and these linkages may be quite subtle or complex. For example, consider the case of a mentally ill person who fires a handgun through the skylight of his apartment because of delusions and hallucinations that space invaders are circling overhead in flying saucers, trying to destroy his brain with invisible radioactive rays that they beam into his home. On charges related to improper or illegal discharge of a firearm, this person may qualify for a judgment of insanity under at least some of the criteria listed in Table I. However, if this person is an ex-convict, then he may also be charged with parole violation because of his being a "felon in possession of a firearm." For this charge, the pertinent inquiry is the nature and quality of his reasoning in relation to obtaining the firearm in the first place. If he obtained the gun many months earlier for reasons unrelated to mental disorder (e.g., "I got it for protection because I live in a bad neighborhood"), presumably the insanity defense would not be relevant to his charge of parole violation.

As should be obvious, the defendant's recollection of his or her thoughts, feelings, and behavior at the time of the offense is critical to mental state at the time of the offense (MSO) evaluations. However, such self-reports may not be completely valid for a variety of reasons. MSO evaluations are often arranged weeks or months after the occurrence of the crime in question or the arrest of the accused. In the intervening period, the defendant's mental condition may have been affected by any number of factors—the natural course of illness, psychiatric treatment received (e.g., in jail or as a result of a finding of incompetence to proceed), or his or her emotional reaction to the offense (e.g., emotional reaction to having killed a loved one). Furthermore, when no other defense is viable, the potential for a defendant to feign insanity in an effort to escape prison must be considered. Consequently, as in the case of other retrospective assessments, the clinical evaluation of MSO often relies extensively on archival, historical, and third-party information.

3.1.4. Competence to Be Executed

In the United States, the courts have held that it is a violation of the Eighth Amendment protection against cruel and unusual punishment for the state to execute an "insane" person. Perhaps the closest the courts have come to defining the criteria for this particular competence was Justice Powell's concurring opinion in *Ford v. Wainwright*, which stated that the state is forbidden from executing only "those who are unaware of the punishment they are about to suffer and why they are to suffer it." A broader criterion suggested by the American Bar Association includes the provision that a defendant be able to recognize and convey to counsel (or a court) any facts that might constitute a basis for appealing this sentence.

Conceptually, these provisions constitute a particular instance of the general criteria for adjudicative competence—the defendant must understand (rationally and factually) the nature and object of the proceedings against him or her and be able to assist counsel. In terms of assessment, the considerations described previously regarding adjudicative competence are relevant here, with a particular substantive focus on the purpose (retribution) and nature (loss of life) of capital punishment.

For a defendant who is judged incompetent for execution, the proceedings are stayed and he or she is referred for psychiatric treatment to restore competence. Perhaps more than for any other forensic issue, mental health professionals feel ethically compromised when they consider that the outcomes of their evaluations or treatment of such defendants may provide the final legal basis for putting them to death. Consequently, this is one area of forensic assessment that some clinicians exclude from their practices.

3.2. Civil Competencies

Apart from the criminal justice system, there are a variety of other social contexts in which concerns about citizens' legal capacities arise. Table II provides a brief description of several such civil competencies.

3.2.1. Guardianship

Some citizens develop mental disabilities that substantially impair their capacity to make, or to communicate, reasonable decisions in day-to-day situations, such as

TABLE II
Civil Competencies

Competence issue	*Fundamental inquiry*	*Legal criteria (generically)*
Need for a guardian to manage one's personal affairs	Should the individual's right to self-determination in day-to-day affairs be limited or removed due to mental incapacity?	Lacking sufficient understanding or capacity to make or communicate responsible decisions (in any of a wide array of day-to-day aspects of living)
Competence to make treatment decisions	Can the person make decisions about medical or psychiatric treatment that are in his or her best interest?	Understand and/or appreciate information relevant to one's illness and alternative treatment options Rationally weigh the risks and benefits, advantages and disadvantages, of treatment alternatives
Competence to consent to research participation	Can the person consent to research participation?	Adequate understanding of the required elements of informed consent in a research consent disclosure Capacity to make a rational decision to participate in the proposed research study
Testamentary capacity	Is the person competent to make a will?	Knowing the extent of their property and the "natural objects of their bounty" (potential heirs) Know that they are making a will and the manner in which their property will be distributed when the will is executed

attending to food and shelter needs, managing finances, and entering contractual agreements (including marriage). One common situation is that of elderly individuals afflicted with Alzheimer's disease or other dementia; some individuals with chronic mental disorders or substance abuse-related illnesses are also so afflicted. Under the parens patria doctrine, most states have provisions for the appointment of responsible third parties (often relatives) to make decisions for individuals who are substantially impaired in this regard.

Of the various legal competencies, the guardianship issue is potentially one of the most sweeping in its scope. As the previous examples illustrate, the ability to "manage one's affairs" can cover a multitude of decisional domains. In recognition of the potentially extensive degree of intrusion that guardianship judgments may entail, the law in many jurisdictions distinguishes between plenary and specific (or limited) guardianships. In the former situation, the court appoints a legal guardian with the authority to make decisions in all (or nearly all) domains of a person's life; in the latter, only specific decisions or types of decisions are placed under the authority of a guardian. A special instance of guardianship is conservatorship, in which a guardian is appointed only for purposes of managing the individual's financial assets (e.g., conducting financial transactions). Out of respect for individual autonomy, there is a preference for specific, as opposed to plenary, guardianship arrangements when the clinical circumstances permit. Thus, when mental health professionals participate in clinical assessments to inform guardianship decisions, their task is not only to determine whether and how severely a respondent's decision making is impaired but also to attempt to determine whether there are areas of decision making in which the respondent's capacity for judgment remains relatively intact.

As with other legal competencies, the mere presence of a mental disorder (i.e., diagnosis) is of little value to the legal decision maker. Although a predicate condition may be required, the legal criteria (depending on, and varying by, jurisdiction) often demand that impairment in specific mental functions (e.g., alertness and attention,

memory, and thought process) be demonstrated and the impact of these impairments on specific decisional behavior (e.g., how to manage finances) be elucidated.

3.2.2. Competence to Make Treatment Decisions

A specific decisional capacity that might be considered a special area of application for guardianship is the capacity to make decisions about medical treatment. This issue arises with respect to both medical and mental health treatment, although it is one for which special procedures and venues (outside formal guardianship proceedings) have developed. As with guardianship proceedings, the situation is one in which a person's right to self-determination (autonomy) is the focal point.

As Table II indicates, the substantive issues involve a person's capacity to comprehend, and to act rationally on, information relevant to making a decision about treatment. Such information includes awareness of one's own illness, the array of potential treatments available, the risks and benefits anticipated with each treatment, the likely course of illness if no treatment is accepted, and so forth. In the context of decision making about whether to accept psychiatric treatment, the courts have recognized varying tests of capacity. These tests parallel, to some degree, the factual understanding, rational understanding, and capacity for autonomous decision making discussed previously in the adjudicative competence context. The person should be aware of the symptoms, nature, and anticipated course of the illness with which he or she is diagnosed; appreciate (i.e., acknowledge) that he or she in fact has the illness and that the probabilities regarding course and outcome—with and without treatment—apply to him or her; and be able to identify and compare risks and benefits in a way that informs judgment about which course of treatment (including refusal of treatment) to follow. Grisso and Appelbaum offer an extensive discussion of the legal framework for treatment decision making and a research-based procedure for clinical evaluations of treatment competence.

Unlike the other legal competencies discussed so far, in some jurisdictions treatment competence determinations are not made in a formal legal arena (i.e., in court) by legal decision makers (i.e., judges or juries). Rather, in some jurisdictions determinations of treatment competence regarding the acceptance or refusal of psychiatric treatment (mainly, psychotropic medications) are left to mental health professionals. Although there may be provisions for administrative review or appeal of clinicians' decisions about the person's need for treatment, the initial judgment whether to impose involuntary treatment may rest with the attending physician and professional colleagues who provide an independent review of that physician's judgment. In these situations, the clinical judgments do involve the judgmental and dispositional components of legal competencies that are left to judges and juries in other contexts.

3.2.3. Competence to Consent to Research Participation

The Nuremburg Code and subsequent guidelines established by governmental administrative bodies, such as the Belmont Report, provide the legal and conceptual frameworks for research involving human subjects. Briefly, no person can be involved as a research participant unless he or she gives competent consent to such participation. In order to give consent, the prospective research participant must be provided information relevant to the decision (whether or not to participate) and demonstrate adequate comprehension of that material. As in the situation involving medical treatment, the research investigator may be liable for a battery against an individual who has not consented to research participation.

As with guardianship and treatment competence, individual autonomy and self-determination are central values at stake in this context. People with sufficient mental capacity to understand research disclosures and the parameters, risks, and potential benefits of research described therein should be allowed to decide for themselves whether or not to participate. However, impaired decision-making capacity can derive from a variety of situational (e.g., stress) or intrapersonal (e.g., mental illness) sources, and some individuals may not be capable of the degree of comprehension and rational judgment necessary to give informed consent. Research suggests that people with dementias (e.g., Alzheimer's patients) and major mental disorders with psychotic features (e.g., schizophrenia) may have the greatest difficulty comprehending research consent disclosures.

Locally, research procedures are approved and monitored by institutional review boards (IRBs). Procedures for eliciting informed consent may vary depending on the type and level of risk that the research entails. Some studies involve minimal or virtually no risk to participants (e.g., an anonymous survey study that does not involve questions about personal or sensitive issues) and in certain circumstances the IRB may waive the requirement that informed consent be obtained. In studies involving greater than minimal risk, investigators are routinely required to obtain written informed consent.

This legal competency is one for which the determination of capacity is not made in a legal forum (e.g., a courtroom) or by a formal legal decision maker (i.e., a judge or jury). Almost universally, the judgment that a person has given competent consent to research participation is made by the study's principal investigator (PI) or research team members trained by the PI in informed consent procedures. The substantive elements of research consent are identified in the U.S. Code of Federal Regulations, and IRBs are responsible to see that PIs include required elements in their written consent disclosures.

There are currently no standardized or objective procedures for evaluating a prospective research participant's level of comprehension of research consent disclosures, which can vary radically in length and complexity from one type of study to another. Miller and Willner recommended brief objective testing of consent disclosure comprehension as one approach to the systematic evaluation of capacity to consent, and Poythress suggested ways that normative assessments of competence to consent might be structured. Currently, research investigators' ad hoc questioning and request for paraphrasing of disclosed material is likely the approach taken with most prospective study participants.

3.2.4. Testamentary Capacity

Planning for the disposition of one's property after death commonly involves making a will. For a will to be valid, the person must have been "of sound mind" at the time that the will was prepared. As Table II indicates, this criterion is ordinarily framed in terms of awareness of the extent of one's property and due consideration of the family members and friends who might be considered the most likely heirs to one's money or estate.

As with other legal competencies, testamentary capacity does not hinge on the mere presence or absence of mental disorder. Mentally disordered individuals may make valid wills, and in order to invalidate a will the courts require a showing that the mental disorder "caused" impairment in the person's ability to evaluate the nature and extent of his or her estate or to consider those to whom the estate might naturally be passed on. Knowledge of one's property does not necessarily have to be detailed, and merely forgetting a distant relative with whom one was marginally involved does not necessarily invalidate a will.

Thus, the assessment will often focus on the individual's capacity to describe the various facets of the estate and to explain the reasons for including, or excluding, potential heirs or the amount of property to be distributed to each. An individual may exclude or severely limit the property planned for distribution to presumably natural heirs (e.g., children or other close relatives) for reasons based in reality (e.g., residual animosity from prior disputes). In contrast, delusional beliefs about such individuals might be judged as a wrongful basis for excluding an otherwise deserving relative or friend, and evidence of such delusions might persuade a court that the person was not of sound mind when the will was prepared.

Challenges to testamentary capacity are likely to be brought by individuals, often family members or friends, concerned about the anticipated distribution of the individual's estate. Such challenges may be brought at the time the person prepares to make a will, in which case the evaluation is a present state assessment. Challenges brought after a will has been made and filed, including challenges after the person is deceased, result in necessarily retrospective evaluations in which the focus is on the testator's mental state at the time the will was made. As in other retrospective evaluations discussed previously, third-party information that helps to reconstruct the testator's thoughts, feelings, and attitudes from the relevant prior point in time may be essential to the assessment.

4. ISSUES IN FORENSIC ASSESSMENT

As should be clear from the foregoing discussion, assessments of legal competencies differ in significant ways from more traditional psychological and psychiatric evaluations that target primarily the diagnosis of mental disorder and the development of treatment plans. The following sections highlight some of the important ethical and assessment features that distinguish evaluations of legal competencies.

4.1. Ethical Issues

Clinicians who participate in the assessment of legal competencies face a number of ethical issues with which they must be prepared to cope. An excellent resource in this regard is the work of Greenberg and Shuman.

One primary issue is that of role clarification. The person whose legal competence is at issue is rarely the person who contacts the clinician for a consultation; rather, the court or an attorney is generally the referring agent, and it is this entity to whom the mental health

professional owes certain obligations. These obligations may need to be discussed and explained in detail with the individual whose competence is being evaluated to ensure that this person does not mistakenly view the clinician as "my doctor" or as having a primarily therapeutic purpose for conducting the assessment. Although there may be some legitimate exceptions, generally mental health professionals should decline to conduct evaluations of legal competency on individuals with whom they have a preexisting therapeutic relationship.

The rules that govern privacy and confidentiality are different in forensic assessments, and these must also be carefully explained to the individual, who may otherwise assume that the usual doctor–patient relationship rules govern subsequent disclosures of information gathered during the assessment.

Another ethical consideration, and one perhaps easily overlooked by forensically naive practitioners, is that of scope of practice. As the foregoing discussion indicates, assessments of legal competence often require an awareness of particular legal tests or criteria, as well as competence in the administration of forensic assessment instruments, that may not be part and parcel of standard mental health training. Thus, some additional training and/or supervised experience in evaluating legal competencies may be required before one can claim adequate education, professional preparation, or competence in conducting such evaluations.

4.2. Instrumentation

It should also be clear from the previous discussion that the assessment approaches and formal instrumentation employed to evaluate legal competencies may vary considerably from those routinely used in therapeutically focused evaluations. Although conventional assessment devices (e.g., EEGs, brain imaging equipment, tests of intelligence or neuropsychological functioning, and personality tests) may have some role in forensic evaluations, legal competencies often involve a more narrow and nuanced inquiry concerning particular functional abilities for which traditional tests were neither developed nor normed. Conventional testing will more likely be helpful in illuminating the background question of an existing mental impairment that might constitute a basis for impaired functioning; independent assessment of the relevant psycholegal abilities and functional capacities is often required and may best be approached using forensic assessments designed especially for such applications. Grisso provides perhaps the most comprehensive review of forensic assessment instruments.

A unique feature of some legal competencies, such as competence to waive Miranda, MSO evaluations, and (some) assessment of testamentary capacity, is their retrospective focus. More often than in therapeutically oriented evaluations, mental health professionals will need to obtain and evaluate archival or third-party information that may be helpful in reconstructing the person's prior mental condition.

4.3. Communicating with Attorneys and Courts

Because evaluation reports from legal competency assessments are written primarily for consumers in the legal community (judges and attorneys), mental health professionals must also learn to write reports that are appropriate for this audience. Perhaps most obvious is the fact that judges and attorneys will not be familiar with the jargon or terms of art that are commonly found in reports written for other mental health colleagues. Clinicians must learn to present clinical material in more of a "plain language" format to ensure that they are understood. Also, as noted previously, clinical diagnoses—diagnostic labels in particular—are less important in forensic contexts. Because of the need to explain whether and how mental impairment affects functional legal capacity, a greater emphasis is needed on description of particular symptomatology and the ways that such symptoms affect competence. Guidelines for communicating with legal professionals can be found in Melton *et al*. and other sources.

See Also the Following Articles

Forensic Mental Health Assessment ■ Psychology and the Law, Overview

Further Reading

Bonnie, R. (1990). Dilemmas in administering the death penalty: Conscientious abstention, professional ethics, and the needs of the legal system. *Law and Human Behavior*, *14*, 67–90.

Committee on Ethical Guidelines for Forensic Psychologists. (1991). Specialty guidelines for forensic psychologists. *Law and Human Behavior*, *15*, 665.

Greenberg, S., & Shuman, D. (1997). Irreconcilable conflict between therapeutic and forensic roles. *Professional Psychology: Research and Practice*, *28*, 50–57.

Grisso, T. (2003). *Evaluating competencies: Forensic assessments and instruments* (2nd ed.). New York: Kluwer/Plenum.

Grisso, T., & Appelbaum, P. S. (1998). *Assessing competence to consent to treatment: A guide for physicians and other health professionals*. New York: Oxford University Press.

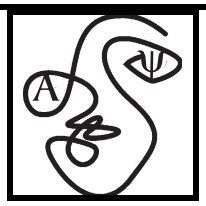
Heilbrun, K. (2001). *Principles of forensic mental health assessment*. New York: Kluwer/Plenum.

Melton, G. B., Petrila, J., Poythress, N. G., & Slobogin, C. (1997). *Psychological evaluations for the courts: A handbook for mental health professionals and lawyers*. New York: Guilford.

Miller, R., & Willner, H. S. (1974). The two-part consent form: Suggestions for promoting free and informed consent. *New England Journal of Medicine*, *290*, 964–966.

Poythress, N. G., Bonnie, R. B., Monahan, J., Otto, R. K., & Hoge, S. K. (2002). *Adjudicative competence: The MacArthur studies*. New York: Kluwer/Plenum.

Rogers, R. (1997). *The clinical assessment of malingering and deception* (2nd ed.). New York: Guilford.



Leisure and Work, Relationship between

Howard E. A. Tinsley and Diane J. Tinsley
Mukilteo, Washington, USA

1. Research on Leisure
2. Emergence of a Psychological Perspective
3. Attributes of Leisure Experience
4. Prerequisites of Leisure Experience
5. Benefits of Leisure Experience
6. Applications of Psychology of Leisure
7. Summary
   Further Reading

GLOSSARY

benefit production process The psychodynamic mechanism that intervenes between a leisure experience and the psychological benefit the individual receives from that experience.

compensation hypothesis A theory that work activities provide only some of the benefits that most people desire, so they engage in leisure activities that provide the benefits their work activities do not provide.

dual-process theories Theories that derive from Wundt's theory of hedonic tone, which postulates that stimulation has both a pleasant and an unpleasant effect, and the individual's evaluation of the stimulus is a function of the relative salience of these pleasant and unpleasant components.

flow A transcendent experience postulated by Csikszentmihalyi to occur as a consequence of an optimal match between the demands of an activity and the skills possessed by the participant.

idle hands theory A widely held belief, rooted in the Protestant work ethnic, that people are happiest and mentally healthiest when they are busy. This view regards industry as good and idleness as bad. Having unoccupied time is seen as psychologically and behaviorally risky.

optimal arousal One of the conditions postulated to be essential for an individual to interpret an experience as leisure. The level of cognitive stimulation (i.e., optimal arousal) that is maximally motivating and satisfying to the brain.

peak experiences A transcendent experience postulated by Maslow that is characterized by high levels of the attributes of leisure experience.

residual/discretionary models Early definitions of leisure that defined leisure in terms of the time remaining after one's obligations had been satisfied.

spillover hypothesis A theory that individuals seek the same benefits from their leisure and work activities.

transcendent leisure experiences theory A theory that when four conditions postulated as necessary to experience leisure are satisfied, (i) an individual will experience the activity as leisure, (ii) the experience will be characterized by high levels of the attributes of leisure experience, and (iii) some psychological needs of the individual will be satisfied.

Psychological investigations of leisure behavior have focused on the causes, attributes, and benefits of leisure experiences. The attributes of leisure experience—the subjective perception of leisure from the perspective of the participant—have received much less attention than the other areas. More attention has been given to the conditions necessary for an individual

©2004 Elsevier Inc.
All rights reserved.

to interpret an experience as leisure (i.e., the causes of leisure experience). Freedom of choice, intrinsic motivation, optimal arousal, and psychological involvement have been proposed as essential conditions for individuals to interpret an experience as leisure. The largest body of research concerns the benefits of leisure experience. Most work on the benefits of leisure has focused on the psychological needs satisfied by leisure experiences, but attention has been given to the effects of leisure experiences on physical health, mental health, life satisfaction, and personal psychological growth. Psychologists also have given attention to the application of leisure psychology to counseling and recreation resource management.

1. RESEARCH ON LEISURE

Only sporadic studies of leisure were undertaken prior to the 1980s. These studies were conducted almost exclusively by psychologists identified with vocational guidance, the forerunner of the modern-day disciplines of vocational psychology and counseling psychology. The most extensive investigation was undertaken under the leadership of E. G. Williamson, a renowned vocational psychologist and one of the intellectual founders of counseling psychology. The influence of vocational psychology is further evident in the narrow focus of this research on leisure activity choice and on the amount of time devoted to various leisure pursuits. A quasi-moral tone is apparent to the contemporary reader in many of these reports; activities such as drinking and playing cards were regarded with less tolerance than is typical in North American society today.

The residual and discretionary definitions of leisure guided leisure research and practice until the 1980s. Both definitions reflect the lay view of leisure as the discretionary time remaining after obligations have been satisfied. Renowned French scholar Joffre Dumazedier gave popular expression to the residual view with his notion of "three eights" (i.e., daily life consists of 8 hours of work, 8 hours of sleep, and 8 hours of leisure). As late as 1980, leisure psychologists, such as Seppo E. Iso-Ahola in his influential book, *The Social Psychology of Leisure and Recreation*, were still defining leisure as the time left after work.

The discretionary model defined leisure as an activity—apart from the obligations of work, family, and society—to which individuals return at will for relaxation, diversion, social participation, broadening their knowledge, or the exercise of their creative capacity. This definition improved on the residual definition by focusing attention on leisure activity rather than time, by recognizing that an individual has family and social obligations, and by suggesting that leisure involves the exercise of free will.

Two newer models of the relation between work and leisure have received research scrutiny. The spillover hypothesis postulates that individuals seek the same benefits from their leisure and work activities. In contrast, the compensation hypothesis postulates that work activities provide only some of the benefits that most people desire, so they engage in leisure activities that provide the benefits their work activities do not provide.

The compensation model makes better intuitive sense. Few people have an occupation that satisfies all of their important desires. It is only logical, therefore, that people obtain the most fulfilling occupation possible and then attempt to satisfy their remaining needs through family, social, and leisure activities. In this instance, however, common sense leads to a false conclusion. The research evidence provides much stronger support for the spillover hypothesis: People generally receive similar benefits from leisure and work activities. This suggests that people seek to obtain the benefits they value most in all aspects of their lives. For example, people who value interpersonal interactions seek both an occupation and leisure activities that provide this benefit.

The first psychologists to study leisure behavior encountered two serious shortcomings in the existing leisure theory and research. First, the residual and discretionary models defined leisure as a void (i.e., time left over after accounting for other activities). These models were not useful in identifying the essential features of leisure, the role of leisure in the life of the individual, or the effects of leisure on the individual. Second, the theory and research on work and leisure through the 1960s reflected a Western European, male, middle-class perspective. The work and leisure experiences of women, working-class individuals, and individuals from other races and cultural heritages were largely ignored. Psychologists have addressed these issues in creating a uniquely psychological perspective on leisure.

2. EMERGENCE OF A PSYCHOLOGICAL PERSPECTIVE

Sporadic investigations of the interaction between individuals and their leisure activities began to appear in the 1960s. The 1974 publication of John Neulinger's

The Psychology of Leisure stimulated the birth of a psychological perspective on leisure and influenced the research and conceptual efforts of psychologists for the next two decades. Neulinger defined leisure as a state of mind and identified perceived freedom and intrinsic motivation as necessary preconditions for leisure to occur. He distinguished among six types of activities based on the extent to which the activity was freely chosen and whether participation was motivated by a desire to receive intrinsic rewards, extrinsic rewards, or a combination of intrinsic and extrinsic rewards. Pure leisure activities occur when individuals freely choose intrinsically satisfying activities; pure job activities occur when individuals respond to behavioral constraints to achieve extrinsic rewards. Although leisure and work were considered to be polar opposites by recreation philosophers, Neulinger argued that leisure was more than the mere absence of work. His definition of leisure and work in terms of the conditions under which they occur enabled scholars to investigate ways in which they are similar.

Initial reactions to Neulinger's model were vigorous. His attempt to redefine leisure as a synonym for "phenomenological experience" was criticized for "confusing communication." His suggestion that only intense experiences could be regarded as leisure ignored the diversity of leisure experiences that occur in everyday life. The ensuing debate focused attention on the nature of leisure experiences and the conditions that are necessary for them to occur. Neulinger's emphasis on the individual's subjective experience and his recognition of the possibility that leisure and work might have similar attributes and benefits highlighted phenomena and issues of interest to psychologists. As a consequence, his model was instrumental in shifting attention away from activity participation patterns and time budget studies.

The theory of transcendent leisure experiences by Howard E. A. Tinsley and Diane J. Tinsley expanded Neulinger's work to include optimal arousal and psychological commitment as additional conditions that are necessary for experiencing leisure. Furthermore, this theory postulated seven attributes of leisure experiences and it suggested that leisure experiences were necessary for physical health, mental health, and personal (psychological) growth.

The transcendent leisure experiences theory highlighted the importance of distinguishing between the causes, attributes, and benefits of leisure experiences. For example, if freedom of choice is a necessary condition for experiencing leisure, individuals who are asked to describe their feelings during a leisure experience might say "I feel free." Since feelings of freedom can be both a necessary prerequisite for leisure experiencing and an attribute of the experience, the failure to distinguish between the causes, attributes, and benefits of leisure experiences leads to confusion. The following sections describe the attributes of leisure experiences, the causes of leisure experiences (i.e., the conditions necessary for leisure experiences to occur), and the benefits of leisure experiences.

3. ATTRIBUTES OF LEISURE EXPERIENCE

Leisure experiences vary from intense, life-transforming experiences to the routine experiences that occur on a daily basis. Nevertheless, psychologists have focused on the most intense forms of leisure experience because the attributes of these experiences are more readily identifiable. Five theorists have postulated attributes of leisure experiences. Abraham Maslow's theory of peak experiences is the earliest and most comprehensive theory. Tinsley and Tinsley's transcendent leisure experiences theory was influenced by Maslow's work, but it deals more directly with leisure behavior. Mihalyi Csikszentmihalyi proposed a theory of flow experiences that formed the basis for Chester F. McDowell's theory of leisure consciousness. Csikszentmihalyi and McDowell focus on the most intense leisure experience. Kathleen D. Noble's theory of transcendent religious and meditation experiences is narrower in scope, but it identifies some attributes that are not included in the other four models. In contrast to the others, Maslow and Tinsley and Tinsley believe that the attributes of routine leisure experiences are the same as but less intense than the attributes of transcendent leisure experiences.

Several common themes are discernable in these theories. The greatest consensus exists for the following attributes of leisure experiences:

Unity: Maslow was the first to propose that the individual and the experience become integrated. Participants report a sense of unification with the world during leisure experiences.

Absorption: Participants focus exclusively on the activity and become unaware of or forgetful of themselves.

Enriched perception: Continued participation across time leads to a greater awareness of the nuances, details, patterns, and complexities of the experience.

Positive outcome: Leisure experiences provide a desirable way of truth finding that can lead to insights into universal truths.

Passive awareness: Participants suspend their will and allow the experience to unfold without attempting to organize or control it.

Loss of inhibition: Anxiety, defensiveness, restraint, and self-criticism are set aside; the individual feels confident, unique, expressive, and daring.

Perceptual distortion: Participants lose track of time and space; hours can seem like minutes.

Feelings of freedom: The individual feels free—unconstrained by rules.

Intrinsic motivation: Individuals gain enjoyment from the mere act of participating in the activity (i.e., from their experiencing of the pattern and structure of the activity and their use of the skills required by the activity).

4. PREREQUISITES OF LEISURE EXPERIENCE

Leisure psychologists have proposed that perceived freedom, intrinsic motivation, optimal arousal, and psychological commitment are necessary for an individual to experience leisure. "Leisure" (and "license") derive from *licere*, a Latin word indicating permission, so it is not surprising that leisure theorists view freedom of choice as essential to experience leisure. Social psychologists refer to freedom of choice as the "illusion of control" because the individual's subjective perception or interpretation of the situation is the critical determinant of the person's experience. Individuals may perceive themselves as freely choosing to engage in behaviors that appear to external observers to be the result of external constraints. The illusion of control can vary from one of no control to complete control.

4.1. Control and Reactance

Psychologists from many subdisciplines have noted the importance of freedom of choice or the illusion of control. Developmental psychologists cite the desire to develop personal competence as a strong influence on the behavior of the child. Motivation psychologists view the desires for personal autonomy and competence as two of the most powerful motivations underlying human behavior. According to social psychologists, the need to feel in control leads people to perceive themselves as having control over situations even when they do not. Illusions have genuine consequences; innovative research studies have revealed that the illusion of control is important for maintaining mental health and successful functioning.

Threatening an individual's sense of freedom generates psychological reactance, an aversive feeling that increases the attractiveness of the threatened freedom, stimulates aggressive feelings, and motivates the person to undertake action to restore the lost or threatened options. Thus, the imposition of constraints has a boomerang effect. The strength of reactance is influenced by the strength of the initial expectation of freedom, the importance of the freedom, the amount of freedom lost, and the effects of the loss of freedom. Leisure psychologists believe that reactance influences people's willingness to accept a substitute activity in place of a lost activity. An individual's willingness to substitute one leisure activity for another is inversely related to the amount of external pressure to make the switch.

4.2. Intrinsic Motivation

Intrinsic motivation refers to the desire for rewards that are obtained from the act of participation. For example, an individual who desires social contact would find visiting with close friends to be intrinsically rewarding because the very act involves social contacts. Extrinsic motivation refers to the desire for rewards that do not occur automatically as a consequence of performing the activity. Common extrinsic motivations include external coercion, the desire to avoid negative outcomes, and rewards such as money, grades, and recognition.

4.2.1. Two-Factor Theory of Motivation

The distinction between intrinsic and extrinsic motivation received attention in vocational psychology long before it became of interest to leisure scholars. The two-factor theory of work motivation views this distinction as critical to understanding vocational behaviors (e.g., job performance) and outcomes (e.g., job satisfaction). Most job satisfaction inventories provide measures of both intrinsic rewards (e.g., feelings of accomplishment and a sense of creativity) and external rewards (e.g., recognition and working conditions). Job-enrichment efforts typically involve changes in the work environment that increase the intrinsic rewards of the job.

4.2.2. Autonomy and Competence

Psychologists who study human motivation believe autonomy, competence, and relatedness are the three

primary intrinsic motives. Autonomy refers to the desire for self-determination and personal choice. As such, autonomy is related to freedom of choice and the illusion of control, and the loss of autonomy is a cause of reactance. Developmental psychologists define competence motivation as the desire to feel in control of the situation and to exercise control for its own sake. These definitions of autonomy and competence reveal a conceptual overlap with the illusion of control. Additional research and conceptual development are needed to illuminate these constructs and eliminate their conceptual redundancy.

4.2.3. Relatedness

Relatedness refers to the desire to feel loved, connected to others, and meaningfully involved with the broader social world. As early as the 1950s, vocational psychologists had identified relatedness as the primary dimension distinguishing among occupations. Early theories postulated that an orientation toward people as opposed to an orientation away from people (or an orientation toward things) was a critical dimension distinguishing among occupations. The strength of a person's preference for working with people is one of the primary factors influencing occupational choices. An extensive body of research supports this view. Some leisure psychologists have also suggested that this is the primary dimension that distinguishes among leisure activities. Research has shown that the desires to affiliate with and nurture others are important dimensions distinguishing among leisure activities.

4.2.4. Overjustification

Providing extrinsic rewards for behaviors that initially are intrinsically rewarding can destroy intrinsic motivation if the extrinsic rewards become too salient (i.e., overjustification). A classic study conducted in 1973 by Lepper *et al*. demonstrated that preschool children who originally chose to play with colorful felt-tip markers during free play were much less willing to play with the markers after an intervening time in which they had been promised and subsequently given a reward for playing with the markers. Leisure theorists believe this is because the extrinsic reward diminished the child's perception of control. They caution against providing extrinsic rewards to encourage individuals (particularly children) to participate in activities that are intrinsically satisfying.

4.2.5. Extrinsic Regulation

Extrinsic rewards regulate the behavior of individuals, thereby lessening their illusion of control. There are three forms of extrinsic regulation. Extrinsic motivation, the most coercive form of extrinsic regulation, occurs when behavior is controlled by external factors. Interjected regulation occurs when internal pressures not associated with one's sense of personal identity (e.g., feeling obligated to perform the behavior) influence behavior. Integrated regulation occurs when behavior is motivated by the person's self-concept (i.e., when the failure to perform the behavior would have negative implications for the person's self-concept).

4.3. Optimal Arousal

Optimal arousal is a third condition believed to be essential to experience leisure. Psychologists believe that individual differences in sensation seeking, curiosity, and exploratory behaviors are the result of individual differences in optimal arousal level.

4.3.1. Homeostasis and Optimal Arousal

Cannon observed in 1932 that all living organisms have a natural tendency to maintain a state of internal balance or equilibrium. He called this tendency homeostasis. The maintenance of homeostasis is largely automatic (e.g., we breathe without having to think about how much oxygen we need). Nissen extended this concept in 1954 to include the function of the brain in perceiving and knowing things. He reasoned that there is an optimal level of cognitive stimulation (i.e., optimal arousal) that is maximally motivating and satisfying to the brain.

4.3.2. Wundt's Law of Hedonic Tone

In 1874, Wundt proposed a general law of hedonic tone to explain how humans interpret stimuli. He theorized that stimulation has both a pleasant and an unpleasant effect. As the strength of a stimulus increases from barely discernable (i.e., neutral) to weak and then to mild, the strength of its pleasantness component increases much more rapidly than the strength of its unpleasantness component. Therefore, the stimulus is perceived as increasingly more pleasant. However, the pleasantness component of each stimulus reaches its maximum while the unpleasantness

component is still relatively weak. Increases in the intensity of the stimulus beyond this point increase only the strength of the unpleasantness component. Wundt's model explains why increases in the intensity of a stimulus (e.g., increasing the tartness of a lemon meringue pie) result in more pleasant sensations up to a point, but further increases beyond that point actually make the sensation less pleasant.

4.3.3. Dual-Process Model

William James, Daniel Berlyne, Marvin Zuckerman, and Charles Spielberger all have used Wundt's dual-process model to explain optimal arousal. Berlyne suggested that increases in the strangeness, novelty, complexity, and information provided by a weak stimulus will produce increases in curiosity or exploratory behavior. As the stimulus becomes stronger, however, its negative component begins to cause avoidance or withdrawal.

Zuckerman offered a similar explanation of sensation-seeking behavior. Research evidence shows that persons who seek varied, novel, or complex sensory experiences tend to need greater levels of stimulation (i.e., have a higher optimal arousal level) than people who do not value these qualities. Research also has revealed the existence of an inverse relation between sensation seeking and anxiety.

Spielberger proposed a dual-process model in which curiosity and exploratory behavior result from the simultaneous influence of curiosity and anxiety. Both are psychological states that arouse and motivate behavior; both fluctuate as a function of novelty, incongruity, and ambiguity. However, these states are antagonistic drives. Pleasure is optimal when there are strong feelings of pleasantness, curiosity, or arousal accompanied by mild feelings of anxiety or fear. As the negative side effects of anxiety and fear increase, the overall quality of the experience becomes less pleasant.

The presence of both curiosity and anxiety in approximately equal strengths leads to uncertainty and an approach–avoidance conflict. Approach–avoidance conflicts occur when a single goal has both undesirable (i.e., risk) and attractive (i.e., possibility of gain) features. These conflicts often produce vacillation. The individual moves closer to the goal until the avoidance motive becomes stronger than the approach motive and then moves away from the goal until the approach motive once again becomes stronger. The result is that the person vacillates around the point at which the approach and avoidance motivations are equally weighted.

4.4. Commitment

A sense of psychological involvement (or commitment) is the fourth condition suggested as a prerequisite for leisure experience. The concept of psychological involvement is useful in distinguishing meaningful participation (e.g., watching one's favorite television show) from inconsequential dalliance (e.g., watching television for 5 minutes while waiting for a companion to arrive). "Commitment" and "involvement" are used somewhat interchangeably in the literature and psychologists have not proposed a formal distinction between them.

In 1947, Muzafer Sherif and Hadley Cantril proposed the concept of psychological involvement to explain the process by which individuals relate their personal sense of identity to a social object (e.g., a sports fan's elation when a favorite team wins). Involvement is a psychological state of arousal, interest, or motivation that has personal relevance to the individual. The strength of involvement influences the extent to which a person will seek out and process information about an activity, expend time and energy pursuing the activity, and continue to pursue the activity across time. Involvement has been investigated most extensively in the area of marketing and advertising behavior.

In leisure theory, involvement (commitment) has been described as a psychological attachment to an activity that guides both long- and short-term behavior. Psychological involvement is determined by a person's satisfaction with the leisure activity, investment in the activity (i.e., the things a person would lose if he or she quit the activity), and perception of the potential benefits available from other activities. A person's level of commitment to a leisure activity predicts his or her activity persistence. Committed individuals have more ability to respond positively to any negative aspects of the activity.

5. BENEFITS OF LEISURE EXPERIENCE

Research has examined a wide variety of potential benefits of leisure, including physiological and physical health benefits, mental health benefits, social benefits, economic benefits, environmental benefits, and benefits to the community. Both individuals and social groups obtain benefits from leisure. Individuals obtain benefits such as increased life satisfaction, but this outcome may also contribute to an improvement in the morale of a family, a classroom, a school, and a community. Thus, social

groups may receive second-order benefits that derive from the first-order benefits individuals obtain from their leisure experiences. Social groups may also receive direct benefits, such as when vandalism is reduced or community pride increased as a result of introducing a new leisure program in the community. Determining the benefits of leisure to social groups is more complicated than determining the benefits received by an individual.

5.1. The Meaning of Benefit

A benefit is a beneficial change or the maintenance of a desirable situation (i.e., a positive outcome). Determining that a positive outcome is a benefit of leisure requires evidence that participation in leisure is responsible for the outcome and that the outcome is an improvement over what would have occurred otherwise. This determination involves a series of subjective judgments. First, psychologists must operationally define leisure. Although much leisure research and theory focuses on the nature of the individual's phenomenological experience, research on leisure benefits has not adopted this approach. Another alternative is to define leisure in terms of a particular setting (e.g., in the mountains) or time (e.g., Sunday afternoons). Research on the benefits of leisure typically has investigated the benefits associated with particular leisure activities (e.g., backpacking or knitting).

Next, the value of the outcome (i.e., positive, neutral, or negative) must be determined. This judgment will vary across individuals and social groups. Some people vandalize property, write graffiti, use drugs, and visit prostitutes during leisure. The individuals involved and some small social groups appear to value the consequences of these activities, but society in general interprets most of the consequences of activities such as these as negative outcomes (i.e., costs). Outcomes such as unwanted pregnancies and contracting a sexually transmitted disease are changed conditions that can occur as a result of leisure activity, but society in general and most individuals would not regard these outcomes as benefits. Leisure psychologists have given virtually no attention to the process by which the value of leisure outcomes is determined.

5.2. Methodological Problems

Finally, it is necessary to determine whether an outcome is a consequence of leisure. A critical difficulty encountered when interpreting research on leisure benefits is that almost all the evidence linking leisure to benefits was obtained using correlational research designs. For example, a typical research finding is that active retirees are more satisfied than passive retirees. Causal inferences cannot be drawn from correlational data.

There are three difficulties inherent in drawing causal inferences from such data. First, the direction of causation is not clear. Perhaps active retirees are more satisfied because they engage in leisure more often, but it is just as plausible to conclude that people who are satisfied with life in general (i.e., not depressed) are more likely to engage in leisure activities. Second, reciprocal causality may exist. Perhaps participation in leisure activities increases retirees' life satisfaction, which causes them to engage in more leisure activities, which in turn further increases their life satisfaction. Another possibility is that a third variable may be the cause. For example, active and inactive retirees may differ in their physical health. Poor health may prevent retirees from participating in leisure activities and detract from their life satisfaction. It should be clear, therefore, that data obtained using correlational research designs do not yield an unambiguous interpretation.

Despite the limitations of correlational research designs, psychologists rely on them almost exclusively when investigating the benefits of leisure. An experimental research design would require psychologists to deprive some individuals of leisure opportunities while enhancing leisure opportunities for others. This would be unethical because of the strong likelihood that leisure has positive, life-enhancing benefits and the possibility that the lack of leisure could harm the individual. Likewise, it would not be ethical for psychologists to manipulate potential leisure benefits, such as life satisfaction, physical health, and morale, to determine if this influenced leisure behavior. For these reasons, it is not possible for psychologists to conduct controlled experiments to identify the benefits of leisure.

5.3. Theories of Benefit Production

Psychologists are unsure of the mechanisms whereby participation in leisure produces the benefits of leisure. Numerous explanations of some aspect of the leisure benefit production process have been offered, but most of these are little more than sketchy outlines of vague ideas. Among these are psychological hedonism theory (psychological well-being requires many small moments of pleasure, which are obtained through leisure), identity affirmation theory (people's interpretation of their own behavior, including their leisure behavior, shapes their personal identities), and buffer and coping theory

(leisure indirectly influences health and well-being by improving one's ability to cope with stressful life experiences).

5.3.1. Idle Hands Theory

A widely held belief is that people are happiest and mentally healthiest when they are busy. This view is rooted in the Protestant work ethnic, which regards industry as good and idleness as bad. Having unoccupied time is seen as psychologically and behaviorally risky. For example, community groups frequently express a concern that the absence of recreation facilities increases the likelihood that adolescents will engage in delinquent behavior. They justify recreation facilities as a means of preventing antisocial behavior.

Governmental agencies also cite the idle hands theory to justify their policies. For example, Daniel L. Dustin's analysis of the recreational pastimes provided at Manzanar, a World War II concentration camp for Japanese American internees, revealed that the camp provided a wide variety of recreational pastimes. Most common were art, followed by sports and games, clubs and organizations, and nature appreciation. These opportunities were plentiful because the war relocation board believed that keeping the internees busy would reduce their resistance and help maintain peace in the concentration camp.

There has been very little research on the idle hands theory. One study of approximately 3000 high school students found that those who were bored with their leisure were more likely to engage in risky behaviors, such as drinking, drug use, bulimia, and attempted suicide. Another study of more than 80 older women found that more active women reported lower levels of guilt, sadness, and emotional distress. Both studies used correlational research designs, however, so unambiguous causal interpretations are not possible for the reasons stated previously.

An even more telling criticism is that idle hands theory does not identify the mechanisms that produce leisure benefits. This view replaces an unknown intervening variable with activity, but it does not explain how activity produces positive consequences nor how low levels of activity produce negative consequences. Further conceptual development is necessary before idle hands theory can be considered to be more than a folk adage stimulated by a belief in the Protestant work ethic. Nevertheless, idle hands theory is widely endorsed by the lay public and government agencies and tacitly assumed by many psychologists.

5.3.2. Theory of Transcendent Leisure Experiences

Tinsley and Tinsley's theory of transcendent leisure experiences integrates need compensation theory, personal growth theory, identity formation and affirmation theory, and buffering and coping theory within the framework of person–environment fit theory. This theory postulates that when the necessary prerequisites are satisfied, participation in leisure activities enables individuals to have leisure experiences. These experiences are characterized by the attributes described earlier and they satisfy some of the individual's psychological needs. For example, intrinsic motivation is a prerequisite for leisure experience, so some of the participant's needs for autonomy, relatedness, and competence will be satisfied whenever a leisure experience occurs. The illusion of control is an attribute of leisure experience, so leisure experiences provide some satisfaction of the individual's needs to feel competent and in control. Leisure experiences are characterized by enriched perception so it is likely that some gratification of the individual's needs for mastery and intellectual stimulation will be obtained. The loss of inhibition that characterizes leisure experiences should reduce an individual's feelings of anxiety, defensiveness, restraint, and self-criticism.

According to the theory of transcendent leisure experiences, the satisfaction of personal needs through leisure experiences enhances an individual's mental health and contributes to improved physical health. These, in turn, increase overall life satisfaction. This process frees psychological resources that otherwise would be needed to satisfy psychosocial needs and maintain health. The result is that the individual has sufficient psychological resources to undertake the self-exploration and personal risks that are necessary for psychological growth. In the terminology of Maslow's theory, leisure experiences enable the individual to strive toward self-actualization.

A goal of Tinsley and associates was to identify the psychological needs satisfied by the full range of leisure activities. Their exhaustive study of psychological needs led to the development of the Paragraphs about Leisure, a research questionnaire that measures the extent to which 44 psychological needs are satisfied by participation in a leisure activity.

Statistical analysis of the data from more than 4000 research participants revealed 11 dimensions that are important in distinguishing among leisure activities. In order of their importance, these dimensions are as follows:

Exertion: Activities high on this dimension emphasize vigorous physical activity; activities low on this dimension offer relaxation and stress reduction. This finding stands in marked contrast to occupational research, which has subsumed this dimension under rubrics such as occupational level and given it scant attention.

Affiliation: Activities high on this dimension provide opportunities to be with other people; activities low on this dimension satisfy the need to be alone and enjoy solitude. The importance of this dimension is consistent with motivation theorists' conclusion that relatedness is one of the primary intrinsic motives and with vocational psychologists' conclusion that orientation toward people versus orientation toward things is the primary dimension underlying vocational interests.

Enhancement: Activities satisfying the need for enhancement give individuals an opportunity to use their talents and develop their skills; activities low on enhancement provide little opportunity for personal development. This and the following dimension are consistent with motivation theorists' conclusion that competence is an important intrinsic motive.

Self-expression: Activities high on self-expression emphasize variety and they afford participants the flexibility to try their own unique approaches and ideas. Activities low on self-expression provide an opportunity to follow a familiar, comfortable routine.

Nurturance: Like affiliation, activities that satisfy this need offer a chance to be with and enjoy the company of others, but they also provide opportunities to support, encourage, and help others. Activities low on nurturance satisfy individuals' needs to pursue their own self-interests.

Compensation: Activities high on compensation provide an opportunity to make up for things that are missing from one's job or other routine activities; activities low on compensation provide pleasures similar to those obtained from other aspects of one's daily routine. As noted previously, the available research provides stronger support for the spillover hypothesis than for the compensation hypothesis.

Sensibility: Activities high on sensibility provide intellectual and artistic stimulation; those low on sensibility provide opportunities to reduce cognitive demands.

Conscientiousness: Leisure activities high on conscientiousness emphasize responsible behavior and personal restraint; activities low on conscientiousness allow individuals to escape feelings of obligation and duty and to focus on pleasant aspects of the present situation.

Status: Activities high on status provide opportunities to receive attention, influence others, and feel important; those low on status allow individuals to avoid attention and the obligations of leadership.

Challenge: Activities high on challenge tend to be goal-directed activities that emphasize personal skill and high levels of performance; activities low on challenge emphasize process more than outcome. Csikszentmihalyi identified challenge as the critical determinant of the flow experience. However, research on the benefits of leisure reveals that leisure activities provide a complex array of psychological benefits and that challenge explains relatively little of the variability among activities.

Hedonism: Activities high on hedonism provide immediate pleasure without the need for long-term planning or commitment; activities low on hedonism require a long-term commitment or special effort and typically involve long-term goals.

6. APPLICATIONS OF PSYCHOLOGY OF LEISURE

6.1. Leisure Counseling

In the 1970s, counseling psychologists and recreation professionals began to advocate the development of a leisure counseling specialty. Papers appeared in the leading journals of both disciplines (e.g., *The Counseling Psychologist* and *Journal of Leisure Research*), and several books appeared that emphasized the need for a leisure counseling specialty, explained the potential benefits of the specialty, and elaborated its conceptual underpinnings.

Advocates of leisure counseling were encouraged by the research evidence that shows a relation between leisure activity participation and enhanced physical and mental health. They reasoned that helping clients select leisure activities that were well suited to meet their needs would result in long-term physical and mental health benefits, and they recommended an integrated approach to career and leisure counseling. Furthermore, they argued that knowledge of the substitutability of leisure pursuits would be useful in helping clients respond to the changes that occur as a result

of factors such as physical injury, unemployment, or the aging process.

Robert T. Overs and associates at the Curative Workshop of Milwaukee, Wisconsin, an agency specializing in rehabilitation counseling, developed an innovative leisure counseling program. Their program was designed to address the needs of clients with disabilities and older persons. They developed a series of questionnaires to be used in assessing the needs of their clients, and they included a research component to evaluate the benefits of their interventions. Their books and monographs represent the most detailed and informative description of an effort to implement a leisure counseling program.

In addition to Overs and associates, several practitioners in the United States began to specialize in leisure counseling. Despite the encouraging results reported by Overs and others, however, mainstream counseling psychology and the recreation profession have given leisure counseling scant attention.

6.2. Recreation Opportunity Spectrum

Beverly L. Driver and associates developed the idea of the recreation opportunities spectrum as a means of summarizing information about the leisure benefits available in an outdoor recreation area. For example, a national park might have many discrete areas. One area might provide opportunities for boating, fishing, hiking, and camping. Another area might provide an opportunity for rock climbing and rappelling, and a third area might be set aside for picnicking and interpretative lectures from park rangers. Each area of the park offers a different sample of leisure benefits because different activities occur in each area.

The recreation opportunities spectrum provides an inventory of the potential benefits provided by an outdoor recreation area. This information is intended for the use of management personnel who oversee public areas such as national parks and wilderness areas. It allows administrators to frame decisions about the addition, modification, or discontinuation of facilities and programs in terms of their effects on the array of psychological benefits provided by the area.

6.3. Amenity Resource Valuation

An issue that often confronts management personnel is how to balance the competing claims of stakeholders to an area. For example, an oil company that wanted to drill for oil in a wilderness area would support its request by promising to add millions of dollars to the economy. When called upon for comment, leisure psychologists and recreation professionals would express concerns that drilling for oil would degrade the quality of the psychological experience for backpackers and other recreational users. In the past, when situations such as this arose the concerns of leisure psychologists and recreation professionals were dismissed summarily. In response to these difficulties, leisure economists and leisure psychologists developed a series of methods for placing a dollar value on leisure experiences.

Three general classes of methods have been developed for valuing leisure amenities. Travel cost methods, the simplest of the approaches, attempt to place a dollar value on recreational experiences by estimating the cost per visit to the site. These methods assume that the user's travel expenditures were made solely for the purpose of using the recreation site. Hedonic pricing methods are based on the notion that the price one would pay for a recreation experience on the open market provides an estimate of the value of an experience obtained in a public area. For example, the cost of renting a canoe from a commercial vendor at a lake in the countryside provides an estimate of the value of canoeing in the boundary waters canoe area. Contingent value methods rely on data obtained using research questionnaires that ask participants how much they are willing to pay for various recreation experiences. Participants' responses are transformed into monetary values that indicate the relative preferences of the respondents for alternative leisure experiences.

7. SUMMARY

In many ways, psychological investigations of leisure have paralleled the study of occupations. This is true not only when viewing the developmental progression of the two disciplines but also when examining the research on the benefits of leisure and work experiences. For example, many of the benefit dimensions identified in research on leisure are similar to those that have been identified in studies of occupational interests, thereby providing a form of convergent validity. However, there is evidence that the relative importance of the benefit dimensions differs. Exertion, a benefit that has received little attention in vocational

psychology, appears to be a more significant dimension than previously recognized.

Important differences are also apparent in the focus of leisure and vocational psychologists. Leisure psychologists have delved into the nature of participants' phenomenological experiences to a greater extent than have vocational psychologists. Although much of the focus has been on extreme (i.e., transcendent) experiences, more is known about the causes and attributes of leisure experiences than is known about work experiences. In contrast, vocational psychology is far more advanced than leisure psychology with regard to the development of practical applications to benefit individuals and social groups.

Almost everyone experiences leisure in some form, and almost everyone enjoys and values leisure. Therefore, leisure represents an important and understudied aspect of life. It seems likely that leisure will be of central importance in the newfound emphasis of the profession on developing a positive psychology of the healthy personality. Leisure behavior provides an amazingly fertile applied laboratory within which to study phenomena of interest to a broad array of subdisciplines in psychology, including cognitive, social, motivation, human factors, vocational, and counseling psychology.

See Also the Following Articles

Competence at Work ■ Intrinsic Motivation and Self-Determination

Further Reading

Crawford, D. W., Jackson, E. L., & Godbey, G. (1991). A hierarchical model of leisure constraints. *Leisure Sciences*, *13*, 309–320.

Driver, B. L., Brown, P. J., & Peterson, G. L. (Eds.). (1991). *Benefits of leisure*. State College, PA: Venture.

Iso-Ahola, S. E. (1980). *The social psychology of leisure and recreation*. Dubuque, IA: Brown.

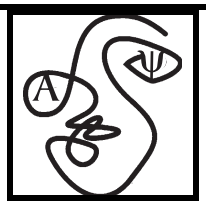
Jackson, E. L., Crawford, D. W., & Godbey, G. (1993). Negotiation of leisure constraints. *Leisure Sciences*, *15*, 1–11.

Neulinger, J. (1974). *The psychology of leisure*. Springfield, IL: Thomas.

Peterson, G. L., Driver, B. L., & Gregory, R. (Eds.). (1988). *Amenity resource valuation: Integrating economics with other disciplines*. State College, PA: Venture.

Tinsley, H. E. A., Hinson, J. A., Tinsley, D. J., & Holt, M. S. (1993). Attributes of leisure and work experiences. *Journal of Counseling Psychology*, *40*, 447–455.

Tinsley, H. E. A., & Tinsley, D. J. (1986). A theory of the attributes, benefits and causes of leisure experience. *Leisure Sciences*, *8*, 1–45.



Lifespan Development and Culture

Patricia M. Greenfield
University of California, Los Angeles, California, USA

Heidi Keller
University of Osnabrueck, Osnabrueck, Germany

Ashley Maynard
University of Hawaii, Manoa, Hawaii, USA

Lalita Suzuki
University of California, Los Angeles, California, USA

1. Conceptual Issues in Cultural Lifespan Psychology
2. Developmental Pathways in a Multicultural Society: Consequences for Application
3. Conclusion
Further Reading

GLOSSARY

apprenticeship A cultural mode of learning where the child (learner) is considered to be a novice who participates in cultural activities as performed by parents (knower) or other experts.

attachment The unique emotional bond that develops between an infant and his or her primary caregivers; the internalization of the mode and quality of these special relationships is supposed to influence the development of later relationships.

collectivism A value system that maximizes and prioritizes harmony and interdependence of the group, especially the family; other terms that are sometimes used to denote this value framework are sociocentrism and communitarianism.

cultural conflict The confrontation of different cultural value systems within an institutional setting (e.g., a school) or a quasi-institutional environment (e.g., a pediatrician's office); in particular, individualist and collectivist value systems have been described as cultural scripts that often specify mutually exclusive choices for individuals' behavior.

cultural learning The processes of verbal interaction, participation in cultural activities, observation, and instruction so as to acquire cultural knowledge.

individualism A value system that maximizes and prioritizes individual choice, independence, and technological mastery.

life stages Functional historical units during ontogeny that can be characterized by universal developmental tasks such as development of attachments during infancy and becoming a productive worker during adulthood (not everyone is a parent); the optimal solution of the universal tasks depends on the sociocultural environment and, therefore, is context and culture specific.

parental ethnotheories Belief systems that specify values with regard to children, parents, and developmental processes; parental ethnotheories are shared among members of cultural communities and set the frame for parental behavior.

peer relations The socioemotional bonds among age mates and between younger children and older children; peers are differentially defined in different cultural environments, being mainly unrelated friends in individualist societies and related children (e.g., siblings, cousins) in collectivist societies.

socialization The process of growing up in a cultural environment; it stresses the active acquisition of information and knowledge of the developing child in the context of a social relational matrix.

©2004 Elsevier Inc.
All rights reserved.

Culture can be perceived as external to the individual in terms of artifacts and institutions; internal to the individual in terms of knowledge, skills, and values; and interactive to the individual in terms of social practices and social interactions. The role of culture in behavior is pervasive and should be considered carefully in any examination of human development over the lifespan. Human development can be understood as the lifelong interactional processes of perceiving and experiencing culture as triggered from the outside, created from the inside, and cocreated with others. Thus, processes of development and processes of culture are inextricably intertwined. Although groups of people share culture and negotiate cultural issues conjointly, at the same time, culture is a subjective individual process.

1. CONCEPTUAL ISSUES IN CULTURAL LIFESPAN PSYCHOLOGY

This section first presents the ecocultural approach as informed by cross-cultural methodology. Then, it presents the sociohistorical approaches that mainly represent the cultural perspective. Finally, it introduces the developmental pathway approach as a new synthesis that incorporates the ecocultural and sociohistorical approaches and takes the whole lifespan into account.

1.1. The Ecocultural Approach

The ecocultural approach emphasizes the role of culture in adapting to the environment and derives from the work of John Berry as well as John and Beatrice Whiting. Berry performed large-scale studies in 21 different sites around the world to analyze how ecocultural information is transformed into styles of perception and cognition. He demonstrated how parameters of the ecological environment (e.g., spatial layout of the landscape) and the mode of subsistence (e.g., farming, foraging) shape the perceptual style of field dependence/independence as the boundedness of perception of patterns to the surrounding visual field.

The Whitings' six-culture study, including research sites from the major parts of the world, was designed on the basis of their ecocultural model linking ecology to modes of subsistence, modes of subsistence to child rearing, child rearing to child personality, and child personality to adult personality The Whitings' study examined this developmental trajectory using a set of consistent methods across different cultures.

Recent theoretical formulations of the ecocultural approach have moved research practice away from the cross-cultural transfer of a set of consistent methods to a more cultural approach with methods appropriate for study in each culture. For example, in 1984, Weisner developed a theory in the Whitings' tradition that considers different aspects of the ecocultural model such as subsistence patterns, child-rearing behavior, and assigned roles for men, women, and children—all of which affect behavior. Subsequently, in 1997, Harkness and Super applied the ecocultural approach to their concept of the developmental niche. The developmental niche is defined as the physical and social settings in which children live, the customs of child rearing, and the psychology of caretakers. Their approach emphasizes the importance of parental cultural belief systems as the link between elements of the larger culture to parenting practices and the organization of the daily routine.

Moreover, the ecocultural approach has become synthesized with evolutionary considerations, specifying the course of development conceptualized in the ecocultural model as following an implicit reproductive logic of optimizing an individual's inclusive genetic fitness resulting from his or her own reproductive success as well as from the reproductive success of genetically related individuals.

1.2. The Sociohistorical Tradition

The sociohistorical research tradition in developmental psychology is derived from the work of Vygotsky. It emphasizes that human development is constructed through social interaction, cultural practices, and the internalization of symbolic cultural tools. An example of a cultural tool that is internalized is money; each culture's system of currency influences the development of certain mental strategies for doing mathematics. Cultural practices and tools are developed over long periods of time. Therefore, they have an important historical dimension.

In addition to artifacts taken in isolation, some cultural psychologists, such as Gauvain and Schliemann and colleagues, have explored the cognitive effects of subsistence practices and trades in terms of representation and cognitive operations, respectively. Cultural practices involve everyday experiences and education (both informal and formal). Therefore, practices such as reading and writing, candy selling, and tailoring shape cognitive development. For example, it has been found that carpenters in Brazil have formal operations in the

domain of proportional reasoning that they use for their trade, even though they do not use this level of reasoning when solving some of Piaget and Inhelder's original tasks. These latter tasks use the much less familiar domains of chemistry and physics.

The locus of social interaction is most often between parents and children due to the importance of parents' caring for their young and for transmitting cultural knowledge from generation to generation. However, teacher–child and peer interactions are also important influences on children's development. Some researchers, such as Maynard and Zukow-Goldring, also recently found that sibling interaction is an important socialization force in both technological and nontechnological societies. Social contexts and the interactional patterns within these social dyads and groups influence both cognitive and social development.

Some researchers, such as Bruner, Tomasello and colleagues, and Shweder, have developed some generalized conceptual approaches to cultural psychology. Bruner noted how the structure of language embodies a structuring of human action in terms of agency and intentional actions. The mutual understanding between people of these categorical structures of intentional action is diagnostic of a theory of other minds. It follows that as language develops, so does this mutual understanding of intentional human action. It is the comprehension of other minds that is central to social sharing and, therefore, to culture.

Tomasello and colleagues, in contrast, emphasized the apprenticeship processes, such as imitation and direct teaching, that enable cultural knowledge to be transmitted from generation to generation. Moreover, they emphasized that cultural learning depends on people's empathic understanding of other minds.

Shweder has developed a notion of environmental constraints versus cultural preferences. This distinction integrates an ecocultural approach with the values approach discussed in the next section. Preferences in cultural practices are guided by an individual's cultural value system. However, the nature of the environment provides constraints to which ideal practices must be adapted.

1.3. Cultural Pathways through Universal Development: A New Synthesis

LeVine was the first to argue that developmental goals guide cultural practices of child care and parenting. He proposed a universal hierarchy of parental goals with survival and health as the foundation, followed by goals relating to economic independence and finally by goals related to the cultural definition of the personality. Systems of developmental goals and values that guide child rearing and socialization for particular parents or groups of parents are called parental ethnotheories of development. Parental ethnotheories can also be thought of as a stage in the lifespan development of cultural values. Two particularly salient models, derived from the industrial and social psychologies of Hofstede, Triandis, and Markus and Kitayama, are individualism and collectivism. These two models are considered to represent different cultural value systems. Each model is adaptive in different ecological environments, thereby incorporating the ecological approach. Each model is actualized through different kinds of social interactions and emphasizes different kinds of social activities and cultural tools, thereby incorporating the sociohistorical approach.

The cultural models are supposed to be linked to the individual's conception of the self in terms of an independent (cf. autonomous, self-contained, and separate) self linked to individualism and an interdependent orientation (cf. allocentrism and relationalism) toward self linked to collectivism. Accordingly, the primary goal of socialization in the independent model is an autonomous, individuated, self-fulfilled person who enters into social relationships and responsibilities by personal choice. The primary goal of socialization in the interdependent model is for the mature person to be embedded in a network of relationships with responsibilities to others. Personal achievements are ideally in the service of a collectivity, most importantly the family. With these cultural goals in mind, many researchers (e.g., Greenfield and Cocking in 1994) have examined the influence of independent and interdependent goals on parenting styles and child development.

The independent pathway can be characterized as follows. Parents and society socialize children to see themselves as predominantly separate from the social context, that is, as bounded individuals who are unitary and stable by definition. The tasks of development are for children to express themselves, their uniqueness, and their own goals. The role of others is to provide reflections for self-evaluation. This model is fostered by, and adapted to, large-scale commercial societies with complex technologies.

The interdependent pathway can be characterized as follows. Parents and society socialize children to

see themselves as connected to the social context and as having flexible and variable selves. The tasks of development are for children to occupy their place in society, to fit in, and to promote the goals of others. The role of others is to provide self-definitions. This model is fostered by, and adapted to, small-scale subsistence societies with simpler technologies.

There are important developmental changes in both of these orientations over the lifespan and over historical time. From a lifespan perspective, the two pathways travel through universal developmental nodes with particular issues that must be negotiated at particular points in each person's lifespan. Each issue presents a particular transformation of the basic value orientations.

For example, a universal node of infancy is attachment, which can be negotiated by maximizing independence (e.g., with a separate crib and baby room as cultural practices) or interdependence (e.g., by cosleeping between mother and baby, by carrying the baby on the caregiver's body). However, Shweder's distinction between preferences and constraints would also call attention to the fact that if a family is living in a small space, ecological constraints might dictate closer sleeping arrangements than their preferred ideal.

Another universal node of development is peer relations. In a culture that prioritizes independence (e.g., mainstream U.S. culture), peer relations focus on separating from the family. In contrast, in cultures that prioritize interdependence, peer relations often focus on relationships with cousins and siblings. Inherent in the pathway approach, with the conception of nodes that organize developmental progress, is the assumption that life trajectories are meaningful, continuous, and coherent. This implies that the early socialization experiences set the stage for the subsequent developmental tasks during the lifespan. After childhood socialization has resulted in an early conception of a more independent or more interdependent self, adolescence constitutes the transition to adulthood. The independent pathway conceives of adolescence as an educational moratorium with the distinct socializing instruction to develop a separate and unique identity. The interdependent pathway conceives of adolescence as a short transition period that is immediately followed by starting one's own family. Accordingly, ages at marriage and birth of the first child differ significantly across these cultural orientations. Adulthood in the interdependent pathway is primarily defined by procreation and parenting. For example, in Cameroonian Nso, a person is considered an adult only if he or she has children (a badge of interdependence), irrespective of age and expertise. Adulthood in the independent pathway is defined, apart from legal boundaries, mainly through educational attainments and economic independence from the family of origin. Finally, old age differs with regard to the greater integration in the families of offspring and especially the respect and care that is focused on the older generation in the interdependence framework. From the historical perspective, Greenfield in 2000 explored the impact of historically driven ecological change on cultural apprenticeship and cognitive representation. Studying two generations of Zinacantec Mayan families over a period of two decades, she demonstrated that the historical transition from a subsistence ecology to a commercial ecology transformed the apprenticeship from a more interdependent style of cultural learning to a more independent one. At the same time, artesanal products, the fruits of a changing apprenticeship process, moved away from a model of community creativity toward a model of individual self-expression.

Critics of these views often argue that the conception of these two cultural orientations is overly simplistic and does not cover the complexity of human psychology within and across cultures. Although everyone exists as both an individual and a group member, and although this dualism is part of the evolutionary heritage and is a structured part of each social group, each social group also denotes a complex system with its own set of priorities. A system of priorities, by its very nature, includes both kinds of components. It is simply that the independence components dominate in one kind of system, whereas the interdependence components dominate in the other. In addition, individualist systems prefer certain forms of social relationships, whereas collectivist systems prefer others. Logically (although less is known about this), individualist systems prefer certain forms of independence, whereas collectivist systems prefer others. Adding to the complexity, priorities shift through processes of ecocultural change and intercultural contact.

Nonetheless, it is heuristic to envision two idealized systems that form two extremes of ecology and values. Each system is not simply a collection of individuals with their individual traits. Instead, each system constitutes a deep structure paradigm that generates many characteristics in a multitude of domains and developmental periods. Although every culture has individual differences, one must note that these differences play a

larger role and are valued more in individualist cultures. Note too that each orientation has its own mode around which individual differences are organized. The two cultural types are structured systems; they are not simply overlapping distributions on a common scale or two. Finally, cultural traits are, by definition, coconstructed and shared among a defined group of people. Consequently, they are less variable than noncultural characteristics.

The nature of development as an individual acquisition process precludes the development of uniformity. Thus, even if there is an ideal modal picture of personality in a cultural community, different real pictures of personality constitute the actual spectrum of the population. Nevertheless, there are obvious differences in developmental pathways that can be captured best by these different value orientations. Indeed, the concept of two paradigmatic pathways allows continuities across developmental pathways to be revealed for the first time. In addition, individuals develop genuine profiles from which they experience themselves as consistent and coherent.

2. DEVELOPMENTAL PATHWAYS IN A MULTICULTURAL SOCIETY: CONSEQUENCES FOR APPLICATION

Intergroup contact is a major source of diverse developmental pathways in multicultural societies such as the United States, Canada, and Australia. As conflicting ethnotheories of development come into contact, many different combinations of values and practices arise. Often, the developmental pathway of the dominant majority tends to be taken for granted as normative, or even universal, rather than as cultural. Intergroup contact raises ethnic identity awareness, often bringing the self-identity of racially distinctive groups to the forefront of developmental issues. This generates more complex processes of cultural learning such as those involved in bilingual and bicultural development.

The distinctive patterns of socialization and development manifest in various groups in the United States and other multicultural societies derive, to a great extent, from ancestral values. Therefore, the specific values that ethnic groups maintain must be understood with respect to ancestral homelands. However, the type of power relation between dominant groups and less dominant groups, as well as the interactions among various groups, also molds cultural development.

Acknowledging cultural influences as a genuine component of development also implies that there is not one (normative) conception of health and psychological well-being. Rather, the definition of healthy development is relative to overarching developmental goals that vary from culture to culture. The following section highlights the implications of incorporating culture into our understanding of developmental problems and health. It introduces two examples. First, it discusses different views on developmental regulation during infancy. Second, it discusses intercultural misunderstandings and their alleviation in the school setting.

2.1. Application of Cultural Pathways During Infancy

As discussed earlier, the early socialization environment reflects cultural values pertaining to highly valued developmental goals within specific contexts. During infancy, developmental contexts and parenting practices that support interdependent developmental goals imply a close union between mother and baby. Cosleeping, nursing on demand, and constant close body contact constitute the invisible foundation of cultural practices with babies. In this context, mothers often meet their infants' needs before such needs are explicitly expressed.

In contrast, in cultures that subscribe to independent developmental goals, developmental contexts and parenting practices during infancy imply separateness and early autonomous functioning. From birth onward, infants are geared to sleep alone and to sleep through the night to establish the circadian rhythm as early as possible. From early infancy, mothers orient their babies to the object world to foster the infants' exploration and to develop attachment to objects that can potentially replace the caregivers (e.g., security blankets, snuggly toys). As a consequence of these different cultural views on closeness and separation, what is regarded as ideal in the interdependent model of the early parent–child relationship (e.g., establishing constant bodily contact between baby and mother) is evaluated as pathological through the cultural lenses of the independent developmental pathway.

For multicultural societies, conflicts between cultural value systems are the inevitable consequence. These conflicts become especially pertinent when both cultural value systems need to interact. For example, the ideology of a society's health care system, including the

expressed views of pediatricians, is usually oriented to the value system of the majority culture.

When the first Turkish families immigrated into Germany during the 1950s, the common practice of Turkish mothers to swaddle their infants for at least a couple of hours per day was strongly discouraged by pediatricians and health care professionals. They claimed that the infants' freedom of movement and their possibilities to explore their environments were largely disrupted by the swaddling; that is, the Germans' independent cultural values did not mirror the interdependent value system of the Turkish immigrants. In this example, the trust of parents in the health care system on which they must rely is deeply affected and constitutes a major risk factor for minority children's development.

Therefore, culturally sensitive developmental counseling and intervention with babies should rely on the assessment of the preferred modes of parenting as related to the cultural values, resources, and constraints that are available to a particular group. Instead of orienting the parents to the cultural ideals of the majority culture (e.g., encouraging working-class Turkish parents to allow their infants more freedom of movement and exploration possibilities), the cultural reality of the parents must be negotiated in terms of meeting the needs of both the infants and the parents.

2.2. Applications of Cultural Pathways in Schools

Another locus of conflict for the interaction of diverse cultural value systems is the school environment, including school–home interaction. Research on Asian American and Latino children in the United States indicates that their families maintain interdependent and collectivist cultural orientations, although the collectivist orientation becomes more restricted in its expression across the generations (e.g., reserved for family celebrations such as weddings). This reduction across the generations results from a contrasting set of priorities in the individualist host culture. These individualist priorities are especially salient in the institution of the school, where they are strongly socialized.

Children and adolescents often experience these divergent value systems as they go between home and school environments in their daily lives. Greenfield, Raeff, and Quiroz analyzed the cultural values of parents, teachers, and children in two schools in Los Angeles. In one study, the sample in one school consisted of groups of predominantly European American children. As expected, the value systems of parents, teachers, and children were consonant with each other (the expressed values were primarily individualist) and there was no cultural conflict among the three groups.

In the other school, the population consisted entirely of Latino immigrant families from Mexico and Central America. As expected, the value systems of parents, teachers, and children differed markedly from each other. The parents held mainly interdependent or sociocentric values, and the teachers held mainly individualist values. The value systems of the children either were between those of the teachers and the parents or were more similar to those of the teachers. Therefore, the children had to serve incompatible developmental goals at the same time (e.g., being obedient and self-effacing at home while being verbally expressive and self-enhancing at school). This type of conflict can produce emotional distress and alienation from either home or school. Although the teachers in both schools were multiethnic, they were monocultural. Regardless of their ethnic backgrounds, their education and teacher training had made them highly individualist in their teaching practices and educational values.

In a subsequent phase of the research, Greenfield, Rothstein-Fisch, and colleagues developed "Bridging Cultures," a series of teacher training workshops to help teachers understand the two developmental pathways and create a bridge between them for students and their families. In this way, it was hoped that the developmental stress caused by the conflict between the two cultural pathways could be ameliorated. Teachers and researchers found that parents participated more in school activities and that children were less conflicted in their classroom interactions.

Another source of developmental stress may emerge from the peer interactions of children with different value systems. In still another study, Suzuki and colleagues analyzed the reactions of adolescent girls to praise and criticism on school sports teams in Los Angeles. These authors demonstrated that reactions to praise and criticism differed with respect to self-enhancement and self-effacement, as predicted by the ethnic origins of the families. Asian American and Latina players valued self-effacing strategies to a greater extent, whereas African American and European American players valued self-enhancing strategies to a greater extent. Self-effacement serves the collectivist goal of fitting into the group, whereas

self-enhancement serves the individualist goal of standing out from the group. In line with their own value preferences, Asian Americans and Latinas generally evaluated self-enhancement negatively (e.g., as conceit), whereas European Americans and African Americans generally viewed self-effacement negatively (e.g., as hurting the team). In sum, there was a tendency to view as negative violations of one's own cultural model. These negative interpretations of differences may form the roots of cross-cultural misunderstanding and conflict.

During a later phase of the project, an intervention was carried out to try to use the same workshop methodology to help teams and coaches understand the two cultural orientations. The goal was to increase intercultural understanding by introducing the notion of distinct basic value systems. Rather than increasing tolerance and understanding for different points of view, the intervention simply made students more collectivist. Further research is needed to develop means appropriate to the high school level to increase tolerance and understanding of values for the two contrasting pathways of development.

3. CONCLUSION

The sociohistorical approach to the role of culture in development emphasizes the importance of social interaction and cultural tools as socializing forces in individual development. The ecocultural approach notes that socialization practices that are typical of a particular culture mold behaviors that are adaptive in particular ecocultural niches. The cultural pathways approach organizes cultural differences into two underlying paradigms that guide universal developmental tasks, such as attachment and peer relations, in distinctly different directions. Each of these pathways is created through cultural processes identified by the sociohistorical approach, and each is adapted to a different type of ecocultural niche. Most important, each pathway provides a cultural lens for viewing and evaluating the social world. Therefore, different cultural frameworks may affect people's relationships with each other, resulting in conflicts in multiethnic societies. Many such conflicts could be avoided if understanding of cultural frames of reference were part of the educational system. In conclusion, the theory of cultural pathways through universal development provides an important theoretical framework for real-world application.

See Also the Following Articles

Child Development and Culture ■ Family and Culture

Further Reading

Berry, J. W., Poortinga, Y. H., Segall, M. H., & Dasen, P. R. (1992). *Cross-cultural psychology*. Cambridge, UK: Cambridge University Press.

Gauvain, M. (1995). Thinking in niches: Sociocultural influences on cognitive development. *Human Development*, *38*, 25–45.

Greenfield, P. M. (2000). Culture and universals: Integrating social and cognitive development. In L. P. Nucci, G. B. Saxe, & E. Turiel (Eds.), *Culture, thought, and development* (pp. 231–277). Mahwah, NJ: Lawrence Erlbaum.

Greenfield, P. M., & Cocking, R. R. (1994). *Cross-cultural roots of minority child development*. Hillsdale, NJ: Lawrence Erlbaum.

Greenfield, P. M., & Suzuki, L. K. (1998). Culture and human development: Implications for parenting, education, pediatrics, and mental health. In I. E. Siegel, & K. A. Renninger (Eds.), *Handbook of child psychology* (5th ed.), Vol. 4: *Child psychology in practice* (pp. 1059–1109). New York: John Wiley.

Harkness, S., & Super, C. M. (1996). *Parents' cultural belief systems: Their origins, expression, and consequences*. New York: Guilford.

Keller, H., & Greenfield, P. M. (2000). History and future of development in cross-cultural psychology. *Journal of Cross-Cultural Psychology*, *31*, 52–62.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.

Maynard, A. E. (2002). Cultural teaching: The development of teaching skills in Maya sibling interactions. *Child Development*, *73*, 969–982.

Munroe, R. L., & Munroe, R. H. (1997). A comparative anthropological perspective. In J. W. Berry, Y. H. Poortinga, & J. Pandey (Eds.), *Handbook of cross-cultural psychology*, Vol. 1: *Theory and method* (pp. 171–213). Boston: Allyn & Bacon.

Ochs, E., & Schieffelin, B. B. (1984). Language acquisition and socialization: Three developmental stories and their implications. In R. Shweder, & R. LeVine (Eds.), *Culture theory: Essays on mind, self, and emotion* (pp. 276–320). Cambridge, UK: Cambridge University Press.

Rogoff, B. (1998). Cognition as a collaborative process. In D. Kuhn, & R. S. Siegler (Eds.), *Handbook of child psychology* (5th ed.), Vol. 2: *Cognition, perception, and language* (pp. 679–744). New York: John Wiley.

Schliemann, A., Carraher, D., & Ceci, S. J. (1997). Everyday cognition. In J. W. Berry, P. R. Dasen, & R. S. Saraswathi (Eds.), *Handbook of cross-cultural psychology*, Vol. 2: *Basic*

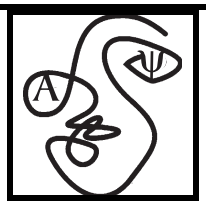
processes and human development (pp. 177–216). Boston: Allyn & Bacon.

Scribner, S., & Cole, M. (1981). *The psychology of literacy*. Cambridge, MA: Harvard University Press.

Super, C. M., & Harkness, S. (1997). The cultural structuring of child development. In J. W. Berry, P. R. Dasen, & J. S. Saraswthi (Eds.), *Handbook of cross-cultural psychology*, Vol. 2: *Basic processes and human development* (pp. 1–39). Boston: Allyn & Bacon.

Weisner, T. S. (1984). Ecocultural niches of middle childhood: A cross-cultural perspective. In W. A. Collins (Ed.), *Development during middle childhood: The years from 6 to 12*. Washington, DC: National Academy Press.

Zukow-Goldring, P. (2002). Sibling caregiving. In M. Bornstein (Ed.), *Handbook of parenting*, Vol. 3: *Status and social conditions of parenting* (pp. 253–286). Mahwah, NJ: Lawrence Erlbaum.



Literacy, Improvement of

Robert C. Calfee

University of California, Riverside, California, USA

1. Literacy: Tools for Thinking and Communicating
2. Improving Literacy: Tasks for Applied Psychology
3. Prospects for Improving Literacy
   Further Reading

GLOSSARY

academic language A formal style of language use found in schools, businesses, and other "ceremonial" settings such as courts and hospitals.

achievement gap The substantial and continuing difference in test performance between students from lower- and middle-class backgrounds, often indexed by "Free and Reduced Lunch" programs.

balanced Programs of reading instruction that aim to combine elements from a variety of more extreme positions.

curriculum-based measurement Assessment systems that attempt to directly measure what is being taught on a daily or weekly basis.

design experiment A research method that combines elements of case studies and field experiments, including both quantitative and qualitative data and emphasizing collaboration between researchers and practitioners.

dynamic assessment Techniques for measuring student learning that allow departures from complete standardization; as with many clinical procedures, the tester uses the informant's responses to decide on additional questions and probes.

dyslexia/learning disabilities Related labels for describing students with reading problems; learning disability is a formal identification used to assign students to special programs supported by federal funding, whereas dyslexia connotes a medical-like syndrome in which brain dysfunction is presumed to be the source of the problem.

initiate–respond–evaluate (IRE) A frequent pattern in classroom discourse during reading lessons whereby the teacher asks a question, a student responds, and the teacher evaluates the response.

linguistic register A style of language use or discourse that is dependent on the individual's background and the demands of a particular situation; in explaining a traffic ticket, a person is likely to use different language when describing the situation to the police officer, to an acquaintance, and to a judge.

No Child Left Behind A multifaceted educational program, promulgated in 2001 by the federal government, calling for widespread testing and accountability.

phonics A system for describing the relations between letter and sound patterns and a set of instructional techniques for teaching students how to decode and spell words; it is often associated with the "simple view" that reading depends on acquiring phonics skills that will then connect to the student's oral language.

student-centered instruction versus direct instruction The contrast between educational techniques that emphasize student input and engagement in a lesson and techniques that emphasize the teacher's role in transmitting information.

whole language A system for early reading instruction that relies on immersion of students in printed material that will then lead to the natural acquisition of literacy skills without the need for extensive phonics instruction.

Since the emergence of psychology as a science, literacy has been a significant topic theoretically, empirically,

©2004 Elsevier Inc.
All rights reserved.

developmentally, and practically. Studies range from word perception to comprehension strategies, from motivational correlates to neurophysiological profiles, and from laboratory studies of college undergraduates to classroom observations of beginning readers. These endeavors often aim to inform basic psychological issues, but possibilities for practical contributions are always close to the surface. "Improving literacy" offers several opportunities for applied psychologists: (a) enhancing performance across the range of tasks in this domain, (b) developing effective methods to support acquisition of skills and knowledge, (c) creating models to undergird instruction and assessment, and (d) incorporating technological inventions that continue to transform the meaning of literacy. In addition, literacy, unlike piloting or driving, is seen as a public "good" that is supported by educational systems and is everyone's business. Psychologists aiming to understand and improve literacy must also deal with social, cultural, organizational, and policy issues. This article is divided into three sections. The first defines literacy for the current analysis, lays out several problems that call for "improvement," and sketches the knowledge base bearing on these problems. The second centers on application and examines the psychological contributions to (a) assessment techniques, (b) evaluation of programs and interventions, and (c) individual differences. The final section offers thoughts about the prospects for improving literacy during the next decade given the current knowledge base and the policy environment.

1. LITERACY: TOOLS FOR THINKING AND COMMUNICATING

1.1. Definition of Literacy

The definition of literacy shapes the problems, claims about what is known about the phenomenon, and strategies for addressing various issues. This section is far more than a philosophical exercise.

Hunt, in a provocative review of DiSessa's *Changing Minds,* suggested in 2003 that "an individual's material intelligence is determined by literacy, the ability to use the intellectual tools that the culture provides." In contrast, reading and writing, viewed as basic skills, may resemble assembly-line routines whereby repetitions with feedback are the primary psychological elements. Hunt's characterization built on the concept of academic language, that is, an acquired and specialized capacity to use various forms of language for thinking and communicating. As basic skill, literacy takes shape as associative learning to be acquired in the primary grades, allowing students to move on to serious topics. As academic register, literacy spans the entire course of schooling and beyond, where print serves as a technology for reducing demands on both short- and long-term memory. The Greeks, before refining their alphabet, employed a variety of mnemonic strategies to support and amplify language capacity, all in the service of organizing thoughts and clarifying communication. This article focuses on the years from preschool through adolescence and spans the continuum from basic to academic literacy.

Academic literacy might seem to entail privilege, and it does. The reference is to a formal style of language use that emphasizes explicitness, expository techniques, coherence, and dependability. The style is typical of educational institutions but also appears in business and the professions. Whether writing a term paper, negotiating a contract, or debating a traffic ticket, individuals in modern society are at an advantage when they can call on the strategies and tools associated with this register.

The acquisition of literacy can be viewed through developmental, sociocultural, and educational lenses. Literacy is rooted in language. Virtually all children progress through developmental stages during which they attain increasing competence in the various elements that constitute language—phonology, semantics, syntax, and discourse. These developments reflect children's sociocultural environment. Much like sponges, children internalize the features of the particular linguistic environments in which they are immersed, leading to variations in linguistic registers or particular styles of use. Children enter schooling with wide variations in their natural registers reflecting their background experiences, but they all possess a full-fledged language system.

A major goal of the educational system is to inculcate students in the academic register. In 1983, Heath offered an exquisite portrait of the encounters of children from varying backgrounds with "school language." The story described the task of learning about print—letters, books, "reading and writing"—but more fundamentally the theme was about variations in language registers and ways of using language individually and in social settings. This article focuses on the role of schools in supporting the acquisition of literacy within a policy context that struggles with numerous issues, including the definition of goals and values.

1.2. Areas Calling for Improvement

What needs to be fixed? At the time of this writing, the answer might seem to be "nearly everything." The U.S. federal government is in the midst of a major initiative, No Child Left Behind, that holds states, districts, and schools accountable for guaranteeing that by 2010 virtually every student will achieve proficiency in the primary school subjects, most especially reading and writing, with dire consequences if the goal is not met. Leading the list of problems are results from the National Assessment of Educational Progress (NAEP) showing that (a) two-thirds of U.S. students are not proficient in reading and (b) there has been little improvement in literacy achievement levels during the past three decades or so. The claim is that the United States is a nation of illiterates, undermining both economic and social capacity.

The second problem is the achievement gap. The NAEP data show that children from poor families do poorly in school, especially in literacy. The differences are enormous and have not diminished during the past several decades despite substantial investments by the federal government through various entitlement programs.

The third problem, which is more diffuse, centers on individual differences in the capacity to read and write and in facility with language. The labels are school centered: special education, learning disability (more than 80% of students in this category have "reading problems"), and dyslexia. The correlates are demographic: students from underrepresented minorities (for whom English is a second language) and boys more often than girls.

The fourth problem is essentially technological on the surface: What is the best method for promoting the acquisition of literacy? The response clearly depends on the definition of literacy, but one might think that for acquisition of basic skills, it should be possible to find an answer. The First Grade Reading Study, a large, federally sponsored field experiment comparing various reading programs, showed some short-term differences favoring intensive phonics programs. But student learning depended more on teachers within programs than on differences among programs. What began as a technological matter of instructional program design metamorphosed into issues of professional development, classroom practice, and organizational interactions.

The context for all of these problems depends on the interpretation of literacy. This article emphasizes behavioral and cognitive perspectives. But in the background are issues of motivation, social and cultural influences, and political–organizational opportunities and constraints.

1.3. The Knowledge Base

During the past several years, the National Academy of Science has conducted several project reports in the areas of language, literacy, and learning to which psychologists have contributed significantly and that provide enormously useful summaries and reviews: Snow and colleagues' 1998 *Preventing Reading Difficulties in Young Children,* the National Reading Panel's 2000 *Teaching Children to Read,* Donovan and Cross's 2002 *Minority Students in Special and Gifted Education,* August and Hakuta's 1997 *Educating Language-Minority Children,* Bransford and colleagues' 2000 *How People Learn,* and Pellegrino and colleagues' 2001 *Knowing What Students Know*. In addition to these resources is the "Rand Report," that is, Snow's 2002 *Reading for Understanding.* Other fundamental resources appear in several handbook series: the three volumes of the *Handbook of Reading Research,* the *Handbook of Research on Teaching Literacy Through the Communicative and Visual Arts,* and the *Handbook of Research on Teaching the English Language Arts* as well as pertinent chapters in the *Handbook of Educational Psychology* and the four volumes of the *Handbook of Research on Teaching.*

The several thousands of pages in these collections constitute a compendium of extraordinary proportions, spanning nearly a half-century of theory and research in the literacy arena, with critical pointers to primary research documents. Included are references to the activities of several educational research centers sponsored by the federal government to explore the application of basic research to school settings as well as numerous large-scale evaluation projects designed to explore the improvement of literacy.

Numerous themes emerge from this collection, three of which are explored in detail in the following section. Three additional points are mentioned briefly due to the relevance to applied psychology. The first point centers on the role of research psychologists as collaborators with other educators, including curriculum developers, assessment designers, and program evaluators. Not so many years ago, experimental and educational psychologists focused on laboratory studies and quantitative data analysis. During the past few decades, increasing numbers of researchers have "taken to the field," moving their investigations into classrooms and schools and joining with other educators in the construction and implementation of programs incorporating theoretical concepts. In many instances, theoretical concepts have not arisen from laboratory findings; rather, field settings have provided the context for generating and refining

theoretical concepts, often taking shape as design experiments that combine the rigor of the laboratory with the messiness of the classroom, joining both quantitative and qualitative sources of data.

A second observation centers on the acquisition of literacy, a process spanning the decade from preschool through adolescence and beyond. The literature provides snapshots of this progression but offers few dynamic portrayals of individual trajectories during the shift from novice to expert. The lack of such intensive data is understandable; research funding comes in 3- to 5-year packages and is "results oriented." The few longitudinal databases available to educators tend to rely on inexpensive and convenient indicators. Much is known about the psychology of skilled reading, that is, the performance of college students in laboratory settings. Perceptual activities in word recognition and sentence comprehension, strategic processes in text comprehension, and "think aloud" protocols during comprehension and composition exemplify domains where substantial work is available regarding expert performance. Less frequent are convincing portraits of the "expert reader/writer" such as those available for mathematicians, aircraft pilots, and chess players. The impediment may be the breadth of the domain encompassed by literacy compared with the other examples. Another set of snapshots comes from the early stages of learning to read, primarily the acquisition of letter–sound skill and knowledge. The bottom line is that research offers perspectives on early and late facets of literacy, but much less is known about the succession of learning in between.

A third point is the challenge of conducting research in school settings, including the task of "scaling up." In other fields of applied psychology (e.g., human engineering, organizational psychology), tradition connects researchers with practitioners. In education, these liaisons are much looser. The researcher approaches schools for "access to subjects," which may be "book-ended" by a technical report of the findings presented to the schools at the end of the project. The client relation, in brief, is often rather distant. During the past several years, the federal government has reintroduced the scaling-up concept, encouraging researchers to move a concept or program from experimentation to wide-scale implementation within a fairly short time frame. The researcher-entrepreneur can be found in other domains, including applied psychology, but in education the tasks are extremely complex, clients range from classroom teachers to state superintendents of education, and researchers might lack the skill, knowledge, and incentives to pursue such endeavors.

2. IMPROVING LITERACY: TASKS FOR APPLIED PSYCHOLOGY

Three arenas have particular relevance in the application of psychological theory and research for the improvement of literacy: assessment, intervention, and individual differences. Assessment is fundamental in operationalizing the constructs that define literacy. Intervention covers a broad and complex array of activities, and psychologists are one of several players in this endeavor. Finally, conceptualizing and investigating individual differences has been a specialization since the beginning of modern psychology.

2.1. Assessment Concepts and Tools

"Literacy is whatever literacy tests measure." This take-off on a well-worn phrase overstates the matter only slightly. Two dimensions frame much of the activity in the assessment field. The first contrast can be characterized by objectives-based versus performance-oriented approaches to capturing information about individuals and groups. The second facet ranges from the classroom-level practice to large-scale tests used for surveys and policy purposes. The two dimensions are not completely independent, but examples can be found in the "four corners" of the matrix.

Across the spectrum, psychological contributions provide the foundation for virtually all assessments. The Standards for Educational and Psychological Testing establish guidelines for policy and practice, laying out fundamental constructs such as validity and reliability, establishing significant practical criteria (e.g., important decisions should rest on multiple sources of evidence), and defining the ethical responsibilities of practitioners.

Within the classroom, assessments can serve several functions such as grading, monitoring progress, and shaping instructional decisions. Assessments may be externally mandated (e.g., the school district requires a writing test), incorporated within instructional materials (e.g., an end-of-unit test), or developed by the teacher. Techniques range from formal activities (e.g., "Close your books and sharpen your pencils; this is a test") to informal and observational (e.g., the teacher listens to a first-grader's reading of a sentence).

Across these settings, the objectives–performance contrast offers an important psychological dimension. In objectives-based assessments, learning outcomes are

defined by highly specific behaviors, and instruments are designed to indicate student mastery of each outcome. In 1971, Bloom and colleagues' *Handbook of Formative and Summative Evaluation of Student Learning* was the progenitor of this approach, which is foundational for the basal reading series widely used for reading instruction in the elementary school grades. The tasks include multiple-choice and short-answer activities, "drawing" answers, and so on. The strategy is to specify outcomes as precisely as possible in behavioral terms. In curriculum-based measurement, a recent addition to this approach, fluency is a primary indicator of mastery (e.g., how many words or sentences can a student read aloud in a minute?).

Performance-oriented techniques rely on cognitive (and metacognitive) constructs as the basis for assessment design built around broader domains and with greater emphasis on explanation. The reading–writing portfolio is a typical artifact in this category, relying on judgmental rubrics to quantify complex performances, much as in the Olympics. A related activity emerges in dynamic assessment, where the task varies in response to student responses so as to explore conditions that evoke student success. This end of the spectrum most often serves for writing assessment.

Externally mandated assessments typically take shape as the standardized multiple-choice test, arguably one of the most significant contributions of applied psychology. Reading is the centerpiece of standardized testing in the elementary grades, soundly established in the "vocabulary-comprehension" subtests. These instruments, although not intended for instructional decisions, can be critical for policy and program evaluation. Writing poses a substantial challenge for large-scale assessment due to the difficulty in capturing complex performances in a cost-effective manner. During recent years, however, several projects have explored the possibilities in this field, most notably in the states of Kentucky, Maryland, Vermont, and California. The importance of these projects arises from showing that it is possible to design systems that connect classrooms with large-scale assessments. Unfortunately, all of these projects have been abandoned in response to the No Child Left Behind initiative.

2.2. Interventions: Programs and People

This segment is organized around three necessarily selective topics: early reading acquisition, the reading–writing connection in the later grades, and pedagogical strategies. The first is a long-standing field of applied psychology, the second illustrates the collaboration of psychologists with others, and the third is more generic.

The First-Grade Reading Study mentioned earlier, Project Literacy, and Chall's landmark review all engaged psychologists in research on beginning reading. Most of the centers and laboratories established by the federal government during the 1960s included projects for the improvement of early reading. Much has been learned as a result of these efforts. For instance, it was previously thought that young children with reading problems suffered from visual perceptual deficiencies, but research demonstrates that the primary correlates are phonological. Another example comes from analyses of English spelling, which is often characterized as inherently chaotic. Research demonstrates the coherence of the system when viewed through historical and phonomorphological constructs. A final example is the discovery that, although direct instruction phonics-oriented programs in the primary grades produce higher scores on standardized tests, (a) program effectiveness varies substantially among teachers within programs and (b) the short-term outcomes tend to "poop out" over the grades. Two general observations seem to hold up. First, much has been learned about the early stages of the acquisition of literacy, including substantive ideas about how to deal with the problems laid out at the beginning of this article. Second, a substantial chasm remains in the practical application of these findings within the complexities of the public education system.

The contributions of psychologists to the study of writing are generally less recognized than those to the study of reading, but the former are nonetheless substantial. The integration of reading and writing, and of language and literacy, becomes increasingly significant as students move through the elementary grades into secondary education and beyond and is not so well studied. From a practical perspective, integration is the key to competence in adult literacy tasks, including college work, jobs in business and industry, and the complexities confronting adults in today's society. Schools tend to teach reading comprehension skills in isolation from composition tasks, and researchers tend to follow this model both conceptually and methodologically. The practical reality is the individual who needs to obtain information from a variety of sources, extract (and record) the key ingredients, and transform the results into a package that serves a practical purpose.

Investigations of pedagogy are generic, although reading often provides the context for research on instructional methods. Psychologists have played a major role in these endeavors, with the first three volumes of the *Handbook of Research on Teaching* being edited by psychologists. Two items warrant particular attention from studies of classroom teaching in the elementary grades: (a) the contrast between strategies that rely on teacher-based direct instruction and those that emphasize student involvement and engagement and (b) explorations of understanding the nature of classroom processes and of the role of teacher decision making in these complex activities.

Probably the most clear-cut finding about classroom instruction is that teachers do most of the talking. This is captured in the initiate–respond–evaluate (IRE) cycle, where the teacher asks a question designed to evoke a particular response, a student responds, and the teacher evaluates the result, turning to other students for a correct answer if necessary. In the primary grades, reading instruction also entails a considerable amount of oral reading, where students take turns reading portions of the day's passage, with IRE events following each student's performance. In the later elementary grades, students are more likely to read the daily passage silently, followed by discussion, but still following the IRE pattern. These techniques, known collectively as direct instruction and typically associated with an emphasis on objectives-based curriculum design, lend themselves to the preprogrammed materials found in teachers' manuals.

The contrastive strategy, often labeled student-centered instruction, places the teacher in the role of building an environment that "scaffolds" student involvement in learning tasks. Students often work in small groups on various projects, with the teacher circulating among the groups to monitor and assist. In whole-language programs, which have gained notoriety during the past decade or so but spring from a much longer history, students can select readings from a range of authentic literature (unlike the tailored passages found in basal readers), with occasional whole-class sessions to introduce learning outcomes (phonics objectives but also vocabulary and comprehension) and to allow students to report their responses to readings. The implicit (and sometimes explicit) assumption is that literacy, like language, will emerge as children are immersed in literary artifacts and supported in gaining mastery of the technical elements, including the technology of print. The assumption that the learning of literacy is as natural as the learning of language is not essential to these programs, nor is it particularly well founded.

Research has illuminated the benefits and limitations of the extreme positions, leading to a search for a middle ground, that is, a balanced approach. Although the concept appears to be attractive, the challenge of defining and designing instruction that incorporates widely dissimilar assumptions and techniques is substantial for both program developers and classroom teachers. It requires more than allocating some time to direct instruction and some time to student-centered activities, a mix that might be programmed into instructional materials and manuals. Investigations of expert teachers show that they create a middle ground that is quite different from either extreme. They typically construct a principled plan for the school year that they constantly modify based on continual assessments of student progress.

Studies of classroom processes and teacher decision making over the past half-century did not begin with expert teachers. During the 1960s, the process–product paradigm employed classroom observations and standardized achievement measures to establish correlational patterns that then led to experimental interventions. For example, if the teacher established effective classroom management techniques, more time was available for academic learning, leading to increased student achievement. By the 1980s, attention turned to issues related to how teachers negotiated the complexities of classroom practice, bringing to bear a range of perceptual, cognitive, and decision-making concepts. Psychologists have played a significant role throughout this progression, moving from a predominantly quantitative perspective to one that incorporates qualitative concepts and techniques and that requires significant collaboration with other disciplines. This paradigm is illustrated by studies of the progression of teachers from novice to professional expertise.

2.3. Individual Differences

Threading through the previous material has been the notion that people vary, an important theme from the earliest days of applied psychology. For the literacy field, attention to individual differences is important in its own right but also provides guidance in practical issues such as identification and assistance for students with special needs. Again, three topics emerge in the literature.

The first and most obvious correlate linking individuals with literacy achievements, in the United States as well as worldwide, is poverty—the achievement gap. Understanding the mechanisms that underlie this correlation, and the exceptions to the generalization, is

perhaps the greatest challenge confronting the United States during the decade ahead, not only for the nation's well-being but also in its role as world leader. Correlational and observational studies of the problem abound. Missing from the mix are sustained and intensive longitudinal studies that might illuminate the processes by which poverty undermines the capacity of students to realize their full potential and to point toward practical interventions. "Success stories" that stand the test of time are in short supply.

The second arena centers around the collection of constructs referred to as learning disabilities or dyslexias. A complete treatment of this topic would require much more space. "Learning disability" is, practically speaking, a bureaucratic label. Students who meet certain behavioral and achievement criteria are identified for federal funding to provide special assistance. The recent history of these programs is generally problematic. Although early intervention seems most likely to make a difference, students often have to "wait to fail" in order to receive assistance. Evaluations suggest that the programs have limited effectiveness. Most of the clients are boys, behavior is the most frequent referral "flag," and low reading achievement is a correlate in more than 80% of the cases.

The various dyslexias spring from medical models, with frequent references to neurological processes and studies of brain functioning. The latter area has burgeoned during recent years with the development of new technologies such as functional magnetic resonance imaging (fMRI). Practically speaking, the overlap with learning disabilities is quite substantial when it comes to the practical aspects of identification and treatment. Substantial federal investments in research and development in these arenas, primarily under auspices of the National Institute for Child Health and Human Development (NICHD), has supported much of the fundamental research in early literacy reported previously, especially in the neurological and brain-imaging areas.

The third area in this category focuses on linguistic and sociocultural differences among students. As noted earlier, language development mirrors children's linguistic environments. Socioeconomic status is a significant dimension in that children from the middle and upper classes are immersed in the academic register virtually from birth. A related issue centers on the increasing proportion of students for whom English is a second language. The growing Hispanic population presents the greatest impact, especially along the nation's southern border, but immigrants from Asia and Eastern Europe are also part of the picture. The issues are quite complex, involving socioeconomic status, cultural styles, home and community languages, and correspondence between student and teacher backgrounds and experiences. English language learners, as they are now labeled, perform poorly on standardized tests across the various programs—bilingual, dual immersion, sheltered English—that have been proposed as remedies. Their performance in mathematics achievement is noticeably stronger, suggesting a linguistic basis for the literacy scores.

3. PROSPECTS FOR IMPROVING LITERACY

Those reading this article are literacy experts who bring their experiences to bear on the definition of the concept, the nature of the problems, and ideas about how to improve literacy. The research literature by no means speaks with a single voice; indeed, the National Reading Panel Report evoked considerable debate due to the selectivity of the investigations and interpretations. Nonetheless, a few generalizations and projections appear to be warranted.

First, several areas call for advancements in theory, research, and practice, most clearly in the achievement gap but also in the pragmatics of curriculum and instruction represented in the concept of balance as an alternative to the continual pendulum swings.

Second, the knowledge base available for improving literacy achievement is substantial, albeit rather contentious, in scope and variety. By any criterion, however, the practices of literacy instruction are quite distant from this base. Literacy teachers learn how to implement various "methods" for teaching reading and writing but are less likely to manifest expertise in linguistics, child development and learning, rhetorical concepts, and methods of behavioral and social science research. The elementary school teacher typically relies on materials and mandates to shape practice along with the occasional counsel of others in the school when time allows, and that is not often.

Third, the application and evaluation of the results of literacy research will call increasingly for studies on application and evaluation, including analysis of the range of policy environments that affect public education. In fields such as engineering and medicine, practitioners are much more closely connected with the methods and findings of the underlying disciplines than with those of education. Once in the classroom,

the individual teacher has considerable leeway about practice (imagine a similar situation for the doctor in the operating room). Schools are described as "decoupled," referring to the disconnections among various levels of the system from district to school to classroom. Those engaged in research and development in the literacy arena are often frustrated by this state of affairs. The remedy might be for these individuals, many of whom are psychologists, to include organizational and contextual elements as an integral part of their research and development agendas.

See Also the Following Articles

Learning ■ Reading, Teaching of ■ Writing, Teaching of

Further Reading

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.

August, D., & Hakuta, K. (1997). *Educating language-minority children*. Washington, DC: National Academy Press.

Barr, R., Kamil, M., Mosenthal, P., & Pearson, P. D. (Eds.). (1990). *Handbook of research on reading* (2nd ed.). New York: Longman.

Barton, P. E. (2003). *Parsing the achievement gap: Baselines for tracking progress*. Princeton NJ: Educational Testing Service.

Berliner, D. C., & Calfee, R. C. (Eds.). (1996). *Handbook of educational psychology*. New York: Macmillan.

Bloom, B. S., Hastings, J. T., & Madaus, G. F. (Eds.). (1971). *Handbook of formative and summative evaluation of student learning*. New York: McGraw–Hill.

Bond, G. L., & Dykstra, R. (1967). The cooperative research program in first-grade reading instruction. *Reading Research Quarterly*, 2(4), 10–141.

Boring, E. G. (1929). *A history of experimental psychology*. New York: Appleton–Century.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn* (rev. ed.). Washington, DC: National Academy Press.

Donovan, M. S., & Cross, C. T. (2002). *Minority students in special and gifted education*. Washington, DC: National Academy Press.

Flood, J., Jensen, J. M., Lapp, D., & Squire, J. R. (Eds.). (1991). *Handbook of research on teaching the English language arts*. New York: Macmillan.

Flood, J., Heath, S. B., & Lapp, D. (Eds.). (1997). *Handbook for literacy educators: Research on teaching the communicative and visual arts*. New York: Macmillan.

Flood, J., Jensen, J. M., Lapp, D., & Squire, J. R. (Eds.). (2002). *Handbook of research on teaching the English language arts* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Foorman, B. R. (Ed.). (2003). *Preventing and remediating reading difficulties: Bringing science to scale*. Baltimore, MD: York Press.

Gage, N. L. (Ed.). *Handbook of research on teaching*. Chicago: Rand McNally.

Heath, S. B. (1983). *Ways with words*. New York: Cambridge University Press.

Hunt, E. (2003). Thinking in the age of the computing machine: Review of A. A. DiSessa, *Changing Minds: Computers, Learning, and Literacy*. *Contemporary Psychology*, *48*, 549–551.

Lidz, C. C., & Elliott, J. G. (Eds.). (2000). *Dynamic assessment: Prevailing models and applications*. Greenwich, CT: JAI.

Linn, R. L. (Ed.). (1988). *Educational measurement* (3rd ed.). New York: Macmillan.

National Assessment of Educational Progress. (2004). *The nation's report card: Reading highlights 2003*. Washington, DC: National Center for Education Statistics.

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature and its implications for reading instruction*. Washington, DC: National Institute of Child Health and Human Development.

Ong, W. J. (2002). *Orality and literacy*. New York: Routledge.

Pearson, P. D. (1984). *Handbook of research in reading*. New York: Longman.

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.). (2001). *Knowing what students know*. Washington, DC: National Academy Press.

Phye, G. D. (Ed.). (1997). *Handbook of academic learning*. San Diego: Academic Press.

Phye, G. D. (Ed.). (1997). *Handbook of classroom assessment*. San Diego: Academic Press.

Shavelson, R. J., & Towne, L. (Eds.). (2002). *Scientific research in education*. Washington, DC: National Academy Press.

Shinn, M. R. (Ed.). (1998). *Advanced applications of curriculum-based measurement*. New York: Guilford.

Snow, C. E. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA: RAND.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Travers, R. M. W. (1973). *Second handbook of research on teaching*. Chicago: Rand McNally.

Wittrock M. C. (Ed.). (1986). *Third handbook on research on teaching*. New York: Macmillan.



Marital Therapy

Angèle Fauchier and Gayla Margolin
University of Southern California, Los Angeles, California, USA

1. Introduction
2. Assessment
3. Common Considerations across Treatment Approaches
4. Specific Therapies
5. Outcome Research
Further Reading

GLOSSARY

behavioral marital therapy Behaviorally based couple treatment focusing on behavior exchange and training in communication and problem-solving skills.

brief strategic couple therapy Family systems-based couple treatment emphasizing a reduction in attempts to solve problems when such attempts may maintain or even worsen problems.

cognitive–behavioral marital therapy Couple treatment combining a behavioral marital therapy approach with a focus on partners' cognitions.

emotion-focused therapy Couple treatment focusing on resolution of insecure attachment patterns and processing of emotion.

insight-oriented marital therapy Couple therapy emphasizing resolution of unconscious maladaptive relationship patterns that arise from previous developmental experiences.

integrative behavioral couple therapy Acceptance-based couple treatment combining a behavioral marital therapy approach with a paradoxical reduction in the struggle to change.

Prevention and Relationship Enhancement Program (PREP) One of several approaches attempting to prevent couple distress; it is often applied to couples prior to their wedding or early in their marriage.

The term "marital therapy" is something of a misnomer because the same therapies apply for couples who are not married, including those who are cohabiting and those who do not live together. However, some treatments specifically target couples at certain points in the marriage process such as prior to marriage or during the first years of marriage, particularly during the transition to parenthood. Often, treatments during the early stages of marriage focus on prevention. Treatments during subsequent stages of marriage usually aim to ameliorate relationship distress. Because individual psychological disorders, such as depression and anxiety, are major risk factors for relationship distress, several investigators have also evaluated the efficacy of treating individual disorders through couple therapy. When relationship problems are contributing to or maintaining individual disorders, couple therapy can assist in treatment of the individual disorder or can change relationship patterns that influence the individual disorder. The current focus, however, is on couple therapies that address relationship problems. Several couple therapies have been empirically supported, meaning that randomized clinical trials have found couples undergoing the treatments to have superior outcomes to couples not experiencing those treatments. Although several approaches have been shown to benefit distressed couples, important questions remain about whether specific therapies can be matched

©2004 Elsevier Inc.
All rights reserved.

to specific couples and whether couple therapy has long-term effectiveness.

1. INTRODUCTION

The vast majority of partners report high satisfaction at the beginnings of their marriages. Yet relationship satisfaction tends to decline over the first 10 years of marriage. Nearly half of all U.S. marriages end in divorce, and most couples are likely to have suffered considerable marital distress prior to divorce. Many other couples whose members remain together also experience marital distress. Marital distress is not stable across the life span. Researchers generally report one of two patterns over time: either a steady decline of marital satisfaction or a complex pattern in which satisfaction is high initially, declines during the child-rearing years, and returns to a higher level during the later years. Marital distress is linked to a variety of individual psychological and physical health problems (e.g., depression, substance abuse). Children growing up in families with marital distress, especially high levels of overt marital conflict, are at risk for psychological problems such as aggression and anxiety. Although approximately 30% of marrying couples receive some form of preventive premarital counseling or relationship education, only 10 to 20% of divorcing couples have met with mental health professionals. The most common sources of disagreement in marriage are money, children and other relatives, the division of housework, communication, and sexual intimacy. It should be noted that the specific interventions available to treat individual or couple sexual problems are not presented in this article.

2. ASSESSMENT

Before commencing treatment, it is important that the clinician perform a thorough assessment of the couple. Such an assessment can encompass a variety of methods. The clinician virtually always interviews the partners, whether separately or together, about their individual and relationship histories, aspects they would like to change, and areas of strength. Couples can complete standardized questionnaires about various aspects of their relationship and individual functioning. These questionnaires provide a measurement of partners' overall marital satisfaction relative to that of other distressed and nondistressed spouses. The questionnaires also evaluate the extent and nature of disagreements in specific relationship domains (e.g., household management, sex, religion, personal habits). Some clinicians perform a formal behavioral observation in which the partners spend a short period interacting with each other so that the clinician can observe their communication style. Patterns of conflict escalation and differences in affective expression are important relationship features that can be observed directly. Clinicians who do not conduct a formal observation invariably observe these dimensions of couple interaction throughout the course of the sessions. Finally, some couples are asked to systematically record information reflecting the nature and tone of their everyday interactions on a daily basis between sessions. These data can include tallies of salient relationship behaviors, both positive (e.g., instances of emotional support) and negative (e.g., arguments), and ratings of important relationship dimensions (e.g., feeling understood by the partner).

3. COMMON CONSIDERATIONS ACROSS TREATMENT APPROACHES

A defining feature of couple therapy is that both partners are seen together in the therapy sessions. This mode of therapy, known as conjoint therapy, raises unique ethical issues. For example, will the therapist keep individual confidences of the partners or will the therapist share information between the partners? How the therapist manages these issues has important consequences for therapeutic alliances and the development of therapeutic goals. Even if the therapist schedules occasional individual sessions with each partner, an overarching principle of conjoint therapy is to balance the therapeutic relationship with both partners and to guard against benefiting one partner to the detriment of the other. Having both spouses in therapy means that the therapist must plan interventions that take into account the well-being and goals of both individuals, and this is a challenge when the partners' goals are at odds.

A common goal in marital therapy, regardless of the specific theoretical perspective, is to improve partners' communication. This goal generally involves exploring new ways of sending and receiving messages so that communications are clearer, less polarized, and more productive. An important benefit of doing conjoint

therapy is that the therapist can directly observe dimensions of the couple's communication style that contribute to relationship distress. The therapist can observe problematic styles such as withdrawal from emotional or conflict-laden discussions and escalation of hostility. The way in which partners communicate in the therapy sessions reveals recurring relationship patterns and provides opportunities to explore how each partner thinks about and responds to those patterns.

Although most marital therapies are designed to relieve distress and to help partners become more satisfied with their relationship, therapy also can be used to help partners decide whether or not to end their relationship. The therapist must balance the extent to which he or she promotes improvement of the partners' relationship versus acknowledges irreconcilable differences between the partners.

4. SPECIFIC THERAPIES

4.1. Behavioral Marital Therapy

Behavioral marital therapy, as originally presented by Jacobson and Margolin, grew out of a social learning perspective of marital distress that focuses on the ways in which partners shape each other's behaviors and on couples' skills deficits. Behavioral marital therapy has two major interventions: behavioral exchange and training in communication and problem-solving skills. Behavioral exchange, through a variety of specific strategies, enables partners to engage in a greater number of pleasing and positive interactions. The therapist assists the partners in planning and carrying out mutually enjoyable activities as well as in showing their affection and caring in small everyday actions. An important part of this intervention is getting partners to switch their focus of attention from negative relationship events to positive ones by becoming more mindful of the exchange of pleasing interactions. The second major intervention is training in communication and problem-solving skills. Communication is one of the most common areas of concern and intervention; partners who communicate well are most likely to maintain high levels of satisfaction. The therapist instructs the partners in listening skills as well as in skills for brainstorming solutions to specific, carefully identified problems. The skills training is designed to help partners discuss important problems in a nonblaming fashion and to develop a specific course of action that reflects the objectives and opinions of both spouses. Concrete plans are made to try the solution on a time-limited basis, to reevaluate the plan at a specific point in time, and (if necessary) to revise and try a modified solution to the problem.

Behavioral marital therapy, with its focus on reducing negative relationship interactions and increasing positive interactions, generally emphasizes the present over the past. One goal of the clinician is to make the partners self-sufficient in solving their own future problems once therapy has ended. Partners who are more attuned to the positive dimensions of their relationship and who have had some success with resolving major conflicts are then more prepared when they are confronted with the future problems that are inevitable in any relationship.

4.2. Cognitive–Behavioral Marital Therapy

Cognitive–behavioral marital therapy, as described by Baucom and Epstein, supplements the elements of behavioral marital therapy by additionally focusing on partners' cognitions about one another and the marriage. Each person brings his or her own interpretation to every moment of every interaction. Sometimes, these interpretations are faulty or can lead to relationship distress as well as individual distress. For example, a wife's disappointment in her husband depends not only on the husband's behavior but also on the wife's expectations of the husband. If the wife's expectations of her husband are unreasonably high, she will find him to be disappointing more often. Cognitive–behavioral therapy might address this problem by focusing on a combination of the wife's expectations and the husband's behavior.

Cognitive–behavioral marital therapy is based on the premise that understanding and changing relationship distress must take account of ways in which partners interpret and evaluate their relationship and one another. In addition to focusing on unrealistic ideals of marriage, this mode of therapy also considers discrepancies in the values and standards that spouses bring to marriage. For example, partners often disagree on how much time should be spent together versus in independent activities, whether it is acceptable to maintain friendships with members of the opposite sex after marriage, how special events (e.g., birthdays, anniversaries) are to be celebrated, and who makes decisions about spending money. In cognitive–behavioral marital therapy, as in behavioral marital

therapy, an important goal is helping spouses to pay attention to the positive aspects of their relationship and not just to notice what is wrong with their relationship. Although this form of therapy also considers partners' behaviors and emotions, its unique contribution is the focus on partners' beliefs, that is, where they came from and what are the advantages and disadvantages of maintaining those beliefs.

4.3. Integrative Behavioral Couple Therapy

Integrative behavioral couple therapy was developed by Jacobson and Christensen to address marital problems that represented serious irreconcilable differences between marital partners and that were not amenable to the change strategies of traditional behavior marital therapy. In integrative behavioral couple therapy, the overriding goal is to increase each partner's acceptance and understanding of the behaviors they would like to see changed in the other partner. Once partners better understand and accept one another and develop a more collaborative approach to their problems, they are more willing to make desired changes. Paradoxically, as each partner "lets go" of the struggle to change the other partner, some changes that could not be achieved before can become possible when both partners accept the situation. Alternatively, for other problems, the therapist's validation of each partner's point of view leads to greater acceptance and greater intimacy, albeit not to specific behavioral change.

An important component of integrative behavioral couple therapy is the formulation of the couple's problems. A couple's problems are understood in terms of thematic issues such as struggles over one partner's desire for closeness and the other's desire for more distance and struggles over which partner has control over a particular area of the relationship. The way in which a couple attempts to solve these differences often polarizes the partners to extreme positions and makes the problems worse. During a feedback session, the therapist shares the formulation about how the partners have become divided over certain issues, with the goal of increasing tolerance of their differences. An important role of the therapist throughout this treatment is to validate each partner's perspective and to highlight how each partner feels hurt and sad, rather than just angry and disappointed, by what has transpired. As initially described in 1981 in Wile's collaborative marital therapy approach, eliciting spouses' vulnerable emotions facilitates closeness and acceptance in their relationship.

4.4. Brief Strategic Couple Therapy

Shoham and Rohrbaugh modified family systems approaches to develop brief strategic couple therapy. The primary focus of this model is the identification and disruption of "ironic processes." Ironic processes are the negative patterns that develop when partners' efforts to solve a problem actually maintain or even worsen the situation. The goal of therapy is to identify and then interrupt this cycle by reducing the partners' attempts to solve the problem. Thus, this approach requires spouses to do less of their currently used solution or even to do the opposite of what they had been doing. An important part of therapy is accepting the partners' definitions of the problem but also reframing and reshaping their understanding of the problem so that they will interact differently around the problem situation. Less of the solution should lead to less of the problem, and this in turn will require even less of the solution, and so on.

Although the focus on interrupting couples' own solutions to their problems is similar to integrative behavioral marital therapy, other features of this approach set it apart from previously described models. In this model, therapy is directed to the specific complaints that clients present. No attempts are made to assess or identify problem areas beyond what the partners themselves have specified. This therapy can be done with one spouse so long as that person is motivated toward change. This approach does not require that the partners understand or have insight into why their problematic patterns have developed. This approach is particularly applicable to those partners for whom skills-based approaches reaffirm and replicate the ineffective solutions that the partners already have tried, for example, structuring more interaction between partners when one spouse already has been criticized about withdrawing from the relationship.

4.5. Emotion-Focused Therapy

As its title suggests, emotion-focused therapy, developed by Johnson and Greenberg, focuses on the emotional experiences of spouses and how these experiences define and maintain interaction patterns. Emotion-focused therapy integrates gestalt/experiential and systems approaches with attachment theory. As an experiential approach, emotion-focused therapy fosters corrective

emotional experiences in therapy sessions. As a systems approach, this model attempts to interrupt repeating negative cycles of interaction. According to this model of therapy, rigid and insecure attachment styles make it difficult for partners to be emotionally open and responsive in marriage, and this in turn creates marital distress. Thus, treatment goals include using the therapy situation to evoke new interactional experiences that create a more secure bond between the partners through reprocessing and restructuring of each partner's inner experiences. When partners develop a sense of security and a better understanding of attachment needs, they can communicate more clearly due to increased empathy and decreased defensiveness.

As contrasted with therapy models that focus on skill building or on relationship beliefs, the emotion-focused therapist helps partners to identify, experience, and express clearer and more poignant emotional needs. These deeper emotional experiences generally reveal vulnerable rather than hostile emotions and lead to a need for connection with the partner. An underlying assumption of this model is that when partners are emotionally engaged with one another, conflict can deescalate and new patterns of interaction can emerge.

4.6. Insight-Oriented Marital Therapy

Insight-oriented marital therapy, developed by Snyder and more recently called affective reconstruction, is similar to emotion-focused therapy in its focus on unidentified and unexpressed affect as the source of relationship distress. However, this approach asserts that maladaptive relationship patterns arise from previous developmental experiences and exist primarily on an unconscious level. Correction of the current marital difficulties is accomplished through the use of clarification and interpretation to uncover each partner's underlying dynamics. A goal of insight-oriented marital therapy is to reconstruct affective experiences and relationship schemas through exploration of each partner's own relationship history. This treatment is flexible and inclusive and often incorporates traditional behavioral techniques of communication training and behavioral exchange.

4.7. Prevention

Prevention-based approaches, such as Markman and colleagues' Prevention and Relationship Enhancement Program (PREP), attempt to improve the quality of the relationship prior to the emergence of problems. Couples can enroll in prevention programs before marriage, early in their marriages, or in long-term committed relationships. Based on research showing that the quality of communication is related to marital distress, skills training in constructive communication, problem solving, and conflict management is a major component of programs such as PREP. Other foci of prevention programs include exploration of each partner's relationship expectations, reinforcement of the partners' commitment, and enhancement of the positive aspects of the relationship. Prevention approaches generally are conducted in groups of couples with a strong educational component through lectures, demonstrations, the practice of new skills through communication exercises, and feedback from the group facilitators. For some couples, the goal of prevention programs is to reduce the likelihood of later distress. For others, the goal is to reverse a downward trend in satisfaction that already has begun.

5. OUTCOME RESEARCH

Most research to evaluate whether couple therapy is effective has compared couples receiving therapy with a control group of couples receiving no therapy. These comparisons have consistently concluded that couple therapy increases satisfaction and decreases distress more than does no treatment. The possibility that another outcome, such as divorce, might be the best outcome for a given couple is generally not addressed in outcome research. The meta-analysis in 1993 by Shadish and colleagues found an effect size of approximately .60, indicating that the average therapy couple has a better outcome than do 70% of controls. Although reviews generally indicate that more than half of couples receiving couple therapy demonstrate reliable improvement, a smaller proportion of couples show improvements by both partners or show movement from distressed to nondistressed status. Long-term follow-up studies have found that most couples maintain gains over the first year but that some couples deteriorate several years following therapy.

Multiple studies have indicated efficacy for the specific types of couple therapy reviewed in this article—behavioral, cognitive–behavioral, integrative behavioral, brief strategic, emotion focused, and insight oriented—as well as for prevention approaches such as PREP. As yet, there is no consistent evidence

that any one of the treatments is better than any other. Although each of these marital therapy treatments seeks to disrupt recurring dysfunctional interaction patterns, the treatments differ in their emphases on behaviors, cognitions, or emotions. Future efforts to match type of treatment to type of presenting problem or to characteristics of the partners might afford higher rates of treatment efficacy.

See Also the Following Articles

Emotion and Culture ■ Psychological Rehabilitation Therapies

Further Reading

Baucom, D. H., & Epstein, N. B. (1990). *Cognitive–behavioral marital therapy*. New York: Brunner/Mazel.

Baucom, D. H., Shoham, V., Mueser, K. T., Daiuto, A. D., & Stickle, T. R. (1998). Empirically supported couple and family interventions for marital distress and adult mental health problems. *Journal of Consulting and Clinical Psychology*, *66*, 53–88.

Christensen, A., & Heavey, C. L. (1999). Interventions for couples. *Annual Review of Psychology*, *50*, 165–190.

Epstein, N. B., & Baucom, D. H. (2002). *Enhanced cognitive–behavioral therapy for couples: A contextual approach*. Washington, DC: American Psychological Association.

Jacobson, N. S., & Christensen, A. (1996). *Acceptance and change in couple therapy: A therapist's guide to transforming relationships*. New York: Norton.

Jacobson, N. S., & Margolin, G. (1979). *Marital therapy: Strategies based on social learning and behavior exchange principles*. New York: Brunner/Mazel.

Johnson, S. M., & Greenberg, L. S. (1995). The emotionally focused approach to problems in adult attachment. In N. S. Jacobson, & A. S. Gurman (Eds.), *Clinical handbook of couple therapy* (2nd ed., pp. 121–141). New York: Guilford.

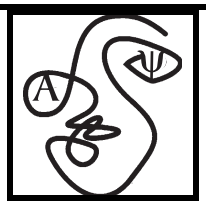
Margolin, G. (1987). Marital therapy: A cognitive–behavioral–affective approach. In N. S. Jacobson (Ed.), *Psychotherapists in clinical practice: Cognitive and behavioral perspectives* (pp. 232–285). New York: Guilford.

Shadish, W. R., Montgomery, L. M., Wilson, P., Wilson, M. R., Bright, I., & Okwumabua, T. (1993). Effects of family and marital psychotherapies: A meta-analysis. *Journal of Consulting and Clinical Psychology*, *61*, 992–1002.

Shoham, V., & Rohrbaugh, M. J. (2002). Brief strategic couple therapy. In A. S. Gurman, & N. S. Jacobson (Eds.), *Clinical handbook of couple therapy* (3rd ed., pp. 5–25). New York: Guilford.

Snyder, D. K., & Schneider, W. J. (2002). Affective reconstruction: A pluralistic, developmental approach. In A. S. Gurman, & N. S. Jacobson (Eds.), *Clinical handbook of couple therapy* (3rd ed., pp. 151–179). New York: Guilford.

Wile, D. B. (1981). *Couples therapy: A nontraditional approach*. New York: John Wiley.



Mathematics, Teaching of

Steven R. Williams
Brigham Young University, Provo, Utah, USA

1. Introduction
2. Current Views of Mathematical Understanding
3. Applications of Cognitive Psychology to Teaching Mathematics
4. The Discursive Turn in Mathematics Education
5. Applications of Cultural Psychology to Teaching Mathematics
6. Conclusion
Further Reading

GLOSSARY

constructivism An epistemological stance claiming that humans actively construct knowledge in ways that are viable and useful for getting along in the world but that knowledge does not necessarily match an external reality. One educational implication of this stance is that knowledge cannot be transferred from person to person, so that teaching becomes a process of providing experiences that will promote and guide knowledge construction.

cultural psychology The study of the role that culture and society play in the mental life of human beings.

understanding Learning with understanding refers to the acquisition of a richly connected network of knowledge, including facts, procedures, concepts, and habits of mind that allow for retrieval and application of knowledge in appropriate situations.

The teaching of mathematics refers to actions specifically intended to promote the learning of mathematics, usually in formal classroom settings. The specific actions taken depend on views of what is fundamental to mathematics itself, what prior knowledge students have, and the environment in which they can best learn mathematical habits of mind.

1. INTRODUCTION

The problem of how mathematics can best be taught and learned has been a matter of concern throughout the past century. Facility with mathematics is seen as a fundamental skill and is increasingly important in modern society. At a time when international studies draw attention to performance gaps, the question of improving students' mathematical understanding has great political and social currency. Applying psychological principles in order to solve this problem seems natural. Indeed, modern psychology has been applied to the learning and teaching of mathematics at least since the 1920s, when E. L. Thorndike published *The Psychology of Arithmetic*. Since that time, various psychological theories, including behaviorism, Piaget's genetic epistemology, and current connectionist and cognitive theories, have been applied with varying success to explain and facilitate mathematical learning.

At the same time, such application is not straightforward. The prototypical question, "How might a teacher best help a student to learn X?" that administrators, policymakers, and even parents would naturally ask presupposes common interpretations of what it means to learn X and in what sense an approach might be "best."

©2004 Elsevier Inc.
All rights reserved.

That these common interpretations might not be in place is partly a result of shifting views of what it means to know mathematics. It is thus important to know how this issue is currently conceived and how it affects the application of psychology to learning mathematics.

2. CURRENT VIEWS OF MATHEMATICAL UNDERSTANDING

In many ways, the foundational issue in mathematics education is the nature of mathematics itself. The decisions we make about teaching mathematics and the kinds of understanding we value in students both depend on what we believe to be most fundamental about mathematics. Mathematics has traditionally been seen as a collection of facts and procedures to be mastered and later applied to appropriate situations. Certainly, learning mathematics does involve mastery of some basic information—multiplication facts, procedures for solving equations, and so forth. Many early applications of psychology to mathematics learning viewed mathematics in just this way. However, it is increasingly recognized that mathematics is more than a collection of facts and procedures. Knowing mathematics involves having the conceptual understanding and habits of mind that enable one to use mathematics powerfully in daily life. Conceptual understanding of mathematics involves understanding how the facts and procedures fit together and when they apply. It involves knowing why procedures work and how to use them to solve problems. In short, it is connected knowledge that provides an organized framework for understanding. Mathematical habits of mind include the ability to reason logically with mathematics ideas, to justify procedures and answers, and to explain mathematical ideas to others—in short, to engage in the practices of mathematics. The acquisition of all these skills, concepts, and habits of mind is what is meant by learning mathematics with understanding. When this becomes our instructional goal, it significantly affects the psychology of teaching mathematics.

3. APPLICATIONS OF COGNITIVE PSYCHOLOGY TO TEACHING MATHEMATICS

Early attempts to apply psychological theory to mathematics teaching were dominated by associationist and behaviorist models of learning. However, the so-called disaster studies of the early 1970s demonstrated that children could attain a degree of computational proficiency without any clear idea of what the computations meant, how they should be applied, or how they were related to other knowledge. Such studies pointed to the inadequacy of "black box" models of human cognition for educational purposes. This realization prompted the more careful study of how children thought (as opposed to whether their answers were right or wrong) and a focus on children's conceptual as well as procedural knowledge. The tools of such study became clinical interviews and the careful observation of children doing mathematics—tools still fundamental to the study of mathematical learning.

As part of this focus on student understanding, and against the behaviorist models, there arose the theoretical stance of constructivism, which has had a profound effect on how mathematics learning is viewed. Grounded in the works of Piaget and explicated by von Glasersfeld and colleagues, constructivism holds that knowledge is actively built up by learners, not in ways that mirror reality but in ways that have viability or fit—ways that help learners make sense of and work in their world. Thus, the understandings built are those that both work for learners and fit with their previous understandings; however, these understandings do not necessarily match an external reality, nor do they necessarily match those of a teacher. One implication is that knowledge cannot be transferred from teacher to student. The teacher can only encourage the construction of viable knowledge by arranging learning situations that support that construction. Teachers who accept constructivist principles focus less on providing students facts and explanations and more on designing the learning environments that help students confront the viability of their knowledge.

Emerging from this background are two critical factors affecting the learning and teaching of mathematics: prior knowledge and mathematical activity. Tasks given to students must connect with students' prior knowledge in order to be approachable at all. Moreover, they must challenge that understanding so there is a need to construct richer and more useful understanding. Thus, tasks must give rise to reflection on what is known and allow for restructuring of knowledge.

3.1. Studies of Early Number Learning

An example of how these ideas bore fruit is provided by the study of how children develop early arithmetic understanding. In the early 1980s, researchers began to document the strategies children used to solve

addition and subtraction word problems as well as the typical developmental paths from simple to more advanced strategies. These studies demonstrated that children's strategies were strongly affected by the action in the word problems (whether two collections of objects were joined together or whether one collection was separated into two, for example) as well as the role of the unknown quantity in the story (e.g., whether the unknown quantity was the result of joining two collections or was one of the two collections being joined).

These studies showed that although children's understanding of arithmetic differed quite dramatically from adults', their knowledge was robust and very useful. Without formal instruction, children were able to develop modeling and counting strategies that allowed for the solution of a wide number of problems. Later studies have repeatedly shown that students come to school with a wide range of mathematical concepts and informal problem-solving strategies. Similarly, studies of older children's mathematics use outside of school have demonstrated that they can develop quite sophisticated and successful contextual strategies for doing mathematics. Research also demonstrates that unless these informal strategies are taken into account when planning instruction, students' abilities to solve problems actually diminish over time. Rather than developing meaningful and useful problem-solving strategies, students come to see mathematics as a collection of arbitrary rules that they have great difficulty applying to word problems.

In contrast, instruction that builds on informal knowledge and provides tasks that push students to reflect on that knowledge has been shown to be effective. For example, studies have shown that elementary teachers who are given information about students' typical strategies, together with tools for assessing students' thinking, can use it to guide their instructional decisions with beneficial results. Thus, a fundamental tool for teachers is the ability to assess their students' understanding of mathematics. The purpose of their assessment is not merely to discern if their students have achieved mastery but also to know the informal knowledge on which students' further understanding can build. Such teaching, heavily dependent on understanding individual students' concepts and strategies, has become an important focus for mathematics educators, with the result that capturing students' mathematical understandings has become a critical issue for both researchers and teachers.

The studies of children's solutions to word problems revealed that the knowledge required to solve word problems did not comprise a set of isolated skills but was indeed a connected network of understandings. These include mathematical ideas associated with number, counting, and order, as well as understanding of problem situations, story structure, and how strategies and problem types were connected. The research also showed that as children developed more advanced solution strategies, their knowledge became even more richly connected and flexible. Thus, "learning with understanding" results in richly connected knowledge structures. Such rich knowledge structures have become the goal of instruction. Because of this, the problem of capturing and describing them becomes critical.

In an attempt to capture students' knowledge, a number of psychological constructs and methods have been employed by mathematics education researchers. These began, perhaps, with Piaget's early structural approach and the approaches of Bruner and Dienes, all of which had a profound effect on early thinking in the field. In the 1980s, information processing psychology became popular as a way to find, describe, and categorize the schemata used by students in the production of mathematical knowledge. Researchers attempted to develop production systems that modeled students' actual performance on mathematical tasks as a means of putting forward rigorous hypotheses about these schemata. Recently, researchers in mathematical thinking have been turning to other areas of cognitive science in order to capture both knowledge structures and the means by which they are created and updated. English and Halford emphasize the importance of analogical processes to learning. Another promising approach is the application of theories of embodied cognition to mathematics learning.

3.2. Summary

The tools of cognitive psychology are critical to the capturing of student's knowledge, which in turn is important to understand both where students are and where we would like them to be. Unfortunately, there has been less success with finding applications of cognitive psychology that help us understand how our students move from current understanding to richer understanding, or of the situations we must create to facilitate this move. It is partly to address this concern that recent work in mathematics education has turned to a focus on social and cultural aspects of learning.

4. THE DISCURSIVE TURN IN MATHEMATICS EDUCATION

Although the focus on individual students' knowledge structures and cognitive processes has met with some

success, it has also been subject to some criticism. In general, cognitive psychology has done a better job of capturing static "snapshots" of students' understanding than in describing the processes of change and learning. Additionally, the cognitive approach often does not capture the full complexity of the classroom, where, after all, much learning and certainly most teaching are thought to occur.

In part because of these criticisms, and in part because of changes in the view of knowledge and learning in general, recent scholarship has focused much more on the social and cultural context of learning mathematics. In some cases, this has taken the form of attending to the ways that social interaction among students and teachers affects the ways that individual understanding is developed in students. A more radical approach calls into question the way we look at mathematical knowledge. From this perspective, understanding mathematics is seen as the ability to participate in particular discourse communities, or communities of practice. This viewpoint has been motivated in part by the recognition that at the highest professional levels of mathematics use, doing mathematics is a social phenomenon with well-established ways to collaborate, share mathematical ideas, and negotiate what counts as mathematically valid. Other influences include the discovery of cultural psychology, as represented in the work of Lev Vygotsky, Michael Cole, and others. All of this has tended to shift attention from mathematics as a purely cognitive activity to include a focus on social practice and cultural acquisition, particularly as played out in classrooms.

In this emerging tradition, the focus shifts from students' individual understanding of mathematical ideas to careful observation of how teachers and students together construct meanings in a social and cultural context, and how these social meanings are appropriated by the students as cognitive tools. The research tools then become participant observation, ethnographic interviews, and similar techniques aimed at describing the social life of the classroom.

5. APPLICATIONS OF CULTURAL PSYCHOLOGY TO TEACHING MATHEMATICS

A good deal of research has focused on learning mathematics in classroom settings, paying particular attention to the sociocultural aspects of learning. In this context, the teachers' responsibilities expand beyond attending to individual students' cognition and providing appropriate mathematical tasks to include the establishment of norms and expectations for social behavior in the classroom. A number of studies have begun to converge on similar descriptions of effective classrooms.

In such classrooms, teachers do provide appropriate tasks that take into account students' prior understanding. However, they also seek to establish expectations that students will (i) seek to learn with understanding, rather than memorizing rules without meaning; (ii) try to understand other students' ideas and solution methods; (iii) explain and justify their thinking to others; (iv) view mistakes as opportunities for learning; and (v) evaluate the usefulness of mathematical ideas based on reason and argument, not solely on the authority of the text or teacher. Such expectations create an environment in which students can safely engage in the mathematical practices of reasoning, justifying, testing conjectures, and communicating and evaluating ideas. In such a classroom, the teacher serves as a representative of the larger mathematical discourse community and can therefore help make needed connections, facilitate the use of established terminology, and provide needed counterexamples or questions as appropriate. In this way, students engage in a sort of "legitimate peripheral participation" in mathematical activity as a way of gaining greater access to the full community of adult users of mathematics. This participation in turn helps students to build the habits of mind that are part of knowing mathematics with true understanding.

6. CONCLUSION

The tools of psychology applied to the problem of teaching mathematics have allowed researchers to gain access to students' existing knowledge, which in turn allows teachers to build on that knowledge through the selection of appropriate tasks. In addition, careful observations of classroom life, made from the perspective of cultural psychology, have allowed teachers to create classrooms in which students engage in mathematical activity as a means of learning what it means to do mathematics.

See Also the Following Articles

Cultural Psychology ■ Learning ■ Teaching Effectiveness

Further Reading

Carpenter, T. P., Franke, M. L., Jacobs, V., & Fennema, E. (1998). A longitudinal study of intervention and understanding in children's multidigit addition and subtraction. *Journal for Research in Mathematics Education*, *29*, 3–20.

English, L. D., & Halford, G. S. (1995). *Mathematics education: Models and processes*. Mahwah, NJ: Erlbaum.

Erlwanger, S. H. (1973). Benny's conception of rules and answers in IPI mathematics. *Journal of Children's Mathematical Behavior*, *1*(2), 7–26.

Fennema, E., & Romberg, T. A. (Eds.). (1999). *Mathematics classrooms that promote understanding*. Mahwah, NJ: Erlbaum.

Hiebert, J., Carpenter, T. P., Fennema, E., Fuson, K. C., Wearne, D., Murray, H., Olivier, A., & Human, P. (1997). *Making sense: Teaching and learning mathematics with understanding*. Portsmouth, NH: Heinemann.

Kilpatrick, J. (1992). A history of research in mathematics education. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 3–38). New York: Macmillan.

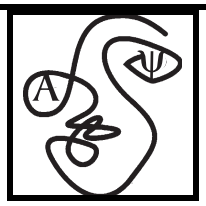
Lakoff, G., & Núñez, R. E. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. New York: Basic Books.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York: Cambridge University Press.

Resnick, L. B., & Ford, W. W. (1981). *The psychology of mathematics for instruction*. Hillsdale, NJ: Erlbaum.

von Glasersfeld, E. (Ed.). (1991). *Radical constructivism in mathematics education*. Dordrecht, The Netherlands: Kluwer.

Wood, T., Nelson, B. S., & Warfield, J. (2001). *Beyond classical pedagogy: Teaching elementary school mathematics*. Mahwah, NJ: Erlbaum.



Measurement and Counseling

Richard M. Lee and Brooke L. Dean
University of Minnesota, Minneapolis, Minnesota, USA

1. Introduction
2. Brief History of Psychological Measurement
3. Relevance of Measurement to Counseling Psychology
4. Counseling Process and Outcome
5. Concluding Remarks
Further Reading

GLOSSARY

assessment The application of measurement to achieve a "snapshot" of a particular variable (e.g., behavior, emotion, ability, interest); counseling psychologists are interested in typical performance on the dimension or variable.
measurement The technique of assigning a dimensional or quantitative representation to a psychological construct; it is usually an attempt to place a numerical value on a person's behaviors, thoughts, emotions, or physiological reactions so as to study a construct.
psychometrics The field of study devoted to psychological measurement; psychometricians investigate the properties of measures used in behavioral science research.
reliability The consistency or dependability of a measuring technique.
validity The extent to which a measurement procedure actually measures the variables intended to be manipulated or measured.

As a cornerstone of counseling psychology, the quantitative method of measurement allows for assessing human behaviors, thoughts, feelings, and attitudes to aid in further developing an individual's potential for personal growth and understanding.

1. INTRODUCTION

Since the age of antiquity, people have sought to differentiate individuals according to abilities so as to select qualified individuals for specific tasks and duties. The earliest record of this type of testing dates back to ancient China (2200 BC) when oral exams were administered to people seeking entrance into the civil service. History is replete with other examples of formal testing on individuals and groups, but these efforts often failed to make reliable and valid predictions. It was not until the 19th century that a breakthrough occurred when psychologists developed quantitative methods to reliably assess individual differences in mental abilities. This quantitative method of assessment, as well as other historical influences that are discussed later, contributed to the emergence of counseling psychology as a specialty field of psychology. In counseling psychology, professionals are trained to assess the interests and abilities of individuals using a variety of methods and to provide reasonable recommendations regarding school, work, and personal life. This article briefly reviews the history of measurement and its relevance to counseling psychology.

©2004 Elsevier Inc.
All rights reserved.

2. BRIEF HISTORY OF PSYCHOLOGICAL MEASUREMENT

The modern-day measurement of psychological attributes dates back to the late 1800s and early 1900s when some of the early founders of psychology, such as Galton, Pearson, Cattell, Binet, Spearman, Thurstone, and Thorndike, attempted to quantify and measure individual differences in mental ability (what later became known as human intelligence). During this heyday period, numerous statistical procedures that now serve as the foundation for psychological measurement were developed. Two of the most important statistical advances were the correlation method and factor analysis. A correlation refers to the strength of the relationship between a pair of variables, with the understanding that knowledge about one variable may carry information about the other variable (e.g., grade point average [GPA] and Scholastic Aptitude Test [SAT]). Factor analysis is a method of analyzing a set of correlations to discover which variables in the set are correlated with each other but are largely independent of the other variables. Thus, factors represent the underlying processes that hold together each subset of correlations.

Using these new statistical procedures, psychologists were able to develop more reliable and valid tests of mental ability or intelligence. These tests were composed of many subtests, and each subtest consisted of multiple items or questions that were intended to measure a specific ability or attribute. The correlation method and factor analysis allowed psychologists to determine whether the items in one subtest (e.g., vocabulary items) were correlated with each other to a greater extent than they were with items from another subtest (e.g., math items). Psychologists also were able to determine which of these tests and subtests correlated with observable behaviors (e.g., school grades and income level).

In short, the pursuit to accurately measure mental ability or intelligence led to the development of the field of psychological measurement, with its emphasis on the construction of reliable and valid scales and instruments, standardization of test administration, and use of probability statistics to aid in the interpretation of test results. Perhaps more important, the advent of psychological measurement as a scientific method allowed psychologists to empirically investigate individual differences in a wide range of attributes and behaviors (not just intelligence) as well as the relationship among these variables. Today, the field of psychological measurement can be divided into psychometrics (i.e., the theory and techniques of measurement) and psychological assessment (i.e., the application of measurement).

The first widespread use of psychological assessment occurred during World War I. Psychologists developed a group intelligence test that was used by the U.S. Army to select and train military personnel for various occupations. This type of assessment began an influx of mass testing and personnel selection—two psychological innovations that continue to play a significant role in society today. Shortly thereafter, psychologists developed measures for a wide range of psychological attributes, including people's attitudes, values, interests, abilities, and personality traits. Among the three most widely known today are the Wechsler Intelligence Scales for children, adolescents, and adults (which measure cognitive functioning); the Minnesota Multiphasic Personality Inventory (which measures personality and psychopathology); and the Strong Interest Inventory (which measures occupational interests).

As society became more concerned about the educational and occupational needs of its citizens, psychologists adapted these measurement advances to everyday practice. The end of World War II is a well-known example of how the fields of measurement and counseling became united. After the war, thousands of veterans needed assistance in their adjustment back to civilian life. They were struggling psychologically with the traumas of combat while seeking higher education opportunities and trying to select appropriate occupations and careers. The Veterans Administration hired counseling psychologists to help returning war veterans overcome personal and vocational adjustment difficulties. Counseling psychologists responded by providing guidance and counseling that was informed by the psychological assessment of individual differences in interests and abilities.

3. RELEVANCE OF MEASUREMENT TO COUNSELING PSYCHOLOGY

Counseling psychology has its earliest historical roots in the vocational guidance movement at the turn of the 20th century. During this period, the rapid industrialization of the United States was accompanied by a diversity of new occupations and careers available to young adults. Vocational guidance helped people

to make appropriate educational and occupational choices based on their abilities and interests. To assist in this guidance process, counseling psychologists turned to recent developments in psychological measurement and the broader study of individual differences to assess these abilities and interests accurately and to ensure a proper fit between people and their environment. The previous example of counseling psychologists working with returning war veterans is an apt case in point and is one of the earliest examples of the use of measurement in counseling psychology.

Measurement and counseling psychology have, in many ways, developed a mutually beneficial relationship over the years. Perhaps the best example of this reciprocity is the development of person–environment fit theory. The person–environment fit theory is derived from the original "trait and factor" work of Frank Parsons, the founder of the vocational guidance movement during the early 1900s. Parsons construed vocational guidance as a three-step process of learning about one's interests and abilities (i.e., trait), understanding the requirements of a given occupation (i.e., factor), and finding a reasonable solution to one's vocational options. E. G. Williamson and John Holland also are credited with advancing the idea that a match or fit between a person's personality and the work environment is the best predictor of occupational satisfaction and success. Although this idea might not seem revolutionary by today's standards, it was groundbreaking during the mid-1900s and set the stage for a plethora of psychological research on the measurement of the person and the environment and on the interaction of person and environment to predict behavior and adjustment.

Furthermore, person–environment fit theory has influenced other developments in measurement and counseling psychology. For example, the concept and measurement of aptitude can be credited to counseling psychologists who sought to understand the readiness and potential of individuals and not simply their levels of intelligence. This emphasis on human potential remains a strong influence in counseling psychology, which as a profession focuses on individual differences in normal development and ways in which to encourage growth and overcome barriers in one's environment. Counseling psychologists also have been at the forefront of multicultural issues in counseling and development. For example, counseling psychologists have contributed to our understanding of test bias in measurement—caused by poorly constructed test items and improper interpretation of test results—and its societal impact on the occupational and educational opportunities and barriers for women and members of minority groups. The ways in which racial/ethnic minorities adapt to a majority culture environment is another major emphasis in counseling psychology. Counseling psychologists have developed numerous theories and measures of racial/ethnic identity development, and have established standards for multicultural competencies in counseling and research, to assess the impact of minority–majority status on behavior and adjustment.

As suggested, numerous psychological instruments were developed and used worldwide in vocational guidance and counseling, including the Strong Vocational Interest Blank (commonly known as the Strong Interest Inventory) in 1943, the American Council on Education Psychology Examination (forerunner to the SAT) in 1943, and the General Aptitude Test Battery (which actually dates to pre-World War II) in 1968. These tests have been revised over the years; however, these tests, along with more recently developed ones, continue to provide individuals with an empirical basis for making informed choices about which occupations and careers to pursue. Counseling psychologists have developed and used psychological instruments for many purposes beyond vocational and educational needs. Some current examples include the measurement of identity development, self-concepts, leisure interests, personal attitudes and biases, cultural values and expectations, and multicultural counseling competencies. In all of these instances, counseling psychology uses measurement to understand the ways in which the person and the environment interact and thereby affect normal development and adjustment.

4. COUNSELING PROCESS AND OUTCOME

By using measurement to assess individual differences in psychological attributes and then applying this knowledge to guidance and counseling, psychologists developed what is now known as differential individual diagnosis. Donald Paterson and others who worked at the University of Minnesota's Employment Stabilization Research Institute during the 1930s are often credited with developing differential individual diagnosis, which involved the use of psychological assessment to form a diagnosis of a client's employment

potential or presenting problem. By the 1970s, this technique was gradually refined from a "shotgun" approach to a "rifle" approach. In the shotgun approach, psychologists administered a standard battery of tests to every client who was referred for an assessment regardless of the presenting problem. In the rifle approach, psychologists selectively administered tests that would provide the specific information necessary to make an appropriate diagnosis and treatment plan. Today, differential individual diagnosis is a cornerstone technique of all counseling approaches.

A nagging question for psychologists has been whether differential individual diagnosis is very accurate. That is, do these diagnoses actually predict behaviors and outcomes? In 1954, Meehl wrote his seminal book, *Clinical Versus Statistical Prediction,* which compared the accuracy of actuarial tables with clinical judgments. To the surprise of many, Meehl found that actuarial (or statistical) method based on psychological test scores was far better than were clinical judgments (or diagnoses) made by psychologists in predicting behaviors and outcomes.

Meehl's finding caused quite a controversy in psychology because it seriously questioned the validity of a psychologist's expertness in making diagnoses. Also, this finding spurred many changes in the ways in which psychologists are trained to use assessment data and make diagnoses. Nowadays, counseling psychology training programs embrace what is referred to as the scientist–practitioner model. This model emphasizes a balanced training in how to conduct counseling and how to conduct empirical research that informs and is informed by the counseling experience. Moreover, the model ensures that psychologists, whether scientists or practitioners, understand how to best formulate, test, and revise hypotheses (or diagnoses) using data generated from psychological assessments and other relevant sources of information (e.g., clinical interviews, medical histories).

Over the years, other questions regarding the effectiveness of counseling have been asked and addressed with the help of refinements in measurement. In his 1966 book, *Revolution in Counseling,* Krumboltz was among the first to ask what is now a popular question in psychology research: Which techniques are most effective with what kinds of clients when used by what kinds of counselors? Counseling researchers have responded to this multifaceted question by focusing mostly on how to best measure and study the process and outcomes of counseling. Researchers have investigated, among many other variables, the ways in which counselors use social influence and persuasion to change client behavior, the importance of a close working relationship between a counselor and a client to effect behavioral changes, and (more recently) the relevance of minority and cultural factors in the counseling experience. In all of these instances, the research has required the construction and validation of scales and instruments that assess client and counselor attributes as well as the interaction between client and counselor.

5. CONCLUDING REMARKS

The role of measurement in counseling psychology helps to distinguish counseling from other helping professions such as psychiatry, social work, and clinical psychology. In contrast to these professions, which tend to operate from a medical model of providing remedial treatment to individuals with more serious psychopathology, counseling psychology uses psychological assessment to identify people's strengths and assets and to evaluate their adjustment to the environment. Armed with these assessment data, counselors are able to provide education, prevention-based programs, and goal-oriented counseling to help individuals prepare and cope with situational problems that arise in everyday living (e.g., school-to-work transition, unemployment, marital and family conflict).

See Also the Following Articles

Attitude Measurement ■ Intelligence Assessment ■ International Classification of Diseases (WHO) ■ Measurement and Counseling ■ Personality Assessment ■ Psychological Assessment, Standards and Guidelines for ■ Psychometric Tests

Further Reading

Dawis, R. V. (1992). The individual differences tradition in counseling psychology. *Journal of Counseling Psychology*, *39*, 7–19.

Gelso, C., & Fretz, B. (2001). *Counseling psychology* (2nd ed.). Belmont, CA: Thomson Wadsworth.

Holland, J. L. (1959). A theory of vocational choice. *Journal of Counseling Psychology*, *6*, 35–45.

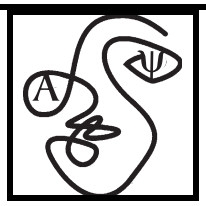
Krumboltz, J. D. (1966). *Revolution in counseling: Implications for behavioral science*. Boston: Houghton Mifflin.

Meehl, P. E. (1954). *Clinical versus statistical prediction*. Minneapolis: University of Minnesota Press.

Sue, D. W., Arredondo, P., & McDavis, R. J. (1992). Multicultural counseling competencies and standards: A call to the profession. *Journal of Counseling and Development*, *70*, 477–486.

Walsh, W. B., & Betz, N. E. (1995). *Tests and assessment* (3rd ed.). Englewood, NJ: Prentice Hall.

Whiteley, J. M. (1984). Counseling psychology: A historical perspective. *The Counseling Psychologist*, *12*, 7–93.



Mental Measurement and Culture

Fons J. R. van de Vijver
Tilburg University, Tilburg, The Netherlands

1. Culture and Mental Measurement
2. Bias
3. Equivalence
4. Cultural Factors in the Assessment of Personality and Intelligence
Further Reading

GLOSSARY

absolutism The viewpoint that holds that cognitive processes are the same across the globe and that emphasizes the context-independent nature of psychological processes.

bias A generic name for potential sources of error in cross-cultural measurement; it refers to the presence of nuisance factors in cross-cultural measurement.

construct bias Bias that occurs when the construct measured is not identical across groups.

equivalence Comparability of test scores obtained in various cultural groups; it involves the question of whether scores obtained in various cultures can be compared in a meaningful way.

item bias A generic term for all disturbances at the item level.

measurement unit equivalence The type of equivalence that instruments show if their scales have the same units of measurement but a different origin (e.g., the Celsius and Kelvin scales in temperature measurement).

method bias A label for all sources of bias emanating from the method and procedure of a study, including factors such as sample incomparability, instrument differences, tester and interviewer effects, and the mode of administration.

relativism The viewpoint that holds that intelligence comes into existence in a specific cultural context and that psychological processes and culture are so closely linked that any attempt to compare these across cultures is futile.

scalar (or full score) equivalence The only type of equivalence that allows for the conclusion that average scores obtained in two cultures are different or equal; it assumes that identical interval or ratio scales are applicable across cultural groups.

structural equivalence When an instrument administered in various cultural groups measures the same construct in all of these groups.

universalism The viewpoint according to which basic cognitive processes are universal, although manifestations of these processes may well vary across cultures.

Comparing individuals from various cultures on psychological constructs, such as depression and intelligence, is more involved than comparing individuals from a single cultural group on the same constructs. This article describes the methodological background of such comparisons. It defines two key concepts: bias and equivalence. Bias is a generic term for potential sources of error in cross-cultural measurement due to the presence of nuisance factors. Equivalence refers to the comparability of test scores obtained in various cultural groups, addressing the question of whether scores obtained in different cultures can be compared in a meaningful way. The article discusses taxonomies of bias and equivalence and uses the assessment of personality and intelligence as an illustration.

©2004 Elsevier Inc.
All rights reserved.

1. CULTURE AND MENTAL MEASUREMENT

Suppose that we are interested in cross-cultural differences in the nature and levels of depression across cultures. Following Berry and colleagues, three approaches can be distinguished: absolutist, relativist, and universalist. A proponent of an absolutist (or etic) position would start from a validated instrument (usually developed in a Western culture) and would administer the instrument in other cultures. Depression is assumed to be a pan-cultural phenomenon, and instruments that were tried and tested in one culture presumably also work in other cultures. Differences in average scores of various countries on the scale are likely to be interpreted as pointing to country differences in depression. The only challenge to this interpretation would be the occurrence of response sets. For example, if there are strict norms in a culture not to disclose information about one's personal feelings to strangers, a depression inventory asking for reports about these feelings may give a poor rendering of the psychological status of the participant or client.

A proponent of a relativist (or emic) view is much less likely to take cross-cultural score differences at face value, as an absolutist does. Depression, like all psychological phenomena, is always expressed in a specific cultural way. Each culture has its own way of dealing with it, and we cannot expect invariance of symptoms and the accuracy of questionnaires developed in another culture. If we want to understand (and measure) depression, we should try to understand the culture and assume the culture's perspective on depression. Instruments are not assumed to travel well across cultures due to the close link between culture and psychological phenomena.

The intermediate view, called universalist, tries to combine both approaches by emphasizing that basic processes are universal but without ruling out the existence of culture-specific expressions of a particular phenomenon. It is an essential property of the universalist position that the (in)variance of behaviors and attitudes associated with an underlying construct (e.g., symptoms of depression) is not assumed but rather is open to empirical examination; cross-cultural studies should determine to what extent depression is pan-cultural, and questions used in Western cultures are relevant for the measurement of depression elsewhere. A universalist attempts to keep an open mind to all possible sources of cross-cultural differences and to verify or falsify specific interpretations of these differences. It is proposed that an exclusive focus on either etic or emic measurement will not solve measurement issues in cross-cultural psychology. Instead, both are needed because they address different aspects of the same phenomenon; an etic measure focuses on the pan-cultural aspects, whereas an emic measure focuses on the culture-specific aspects. A universalist approach may offer the most fruitful approach to cross-cultural differences because it is the only approach that is sufficiently flexible to deal with the multiple ways in which cultural factors can affect mental measurement.

A good starting point for appreciating the universalist approach is provided by an overview of bias, a generic name for potential sources of error in cross-cultural measurement. The next two sections of the article discuss bias and equivalence (comparability of test scores); an overview of the concepts is provided in Table I. The subsequent section gives a brief overview of the main findings in the cross-cultural studies of personality and intelligence. Conclusions are drawn in the final section.

2. BIAS

Bias refers to the presence of nuisance factors in cross-cultural measurement. If scores are biased, their psychological meaning is dependent on the culture in which they have been obtained. Differences between cultural groups in assessment outcome are due, at least to some extent, to cultural or measurement artifacts. For example, differences in scores on self-esteem obtained by Korean and American participants may be influenced by a norm of modesty, which is stronger in Korea than in the United States, and may be contaminated by social desirability.

2.1. Construct Bias

Construct bias occurs when the construct measured is not identical across groups; examples of sources of bias can be found in Table II. Construct bias precludes the cross-cultural measurement of a construct with the same measure. An example can be found in work on filial piety, defined as the psychological characteristics associated with being perceived as a good son or daughter. The Chinese conception, according to which children are supposed to assume the role of caretaker of elderly parents, is broader than the Western concept, which is more restricted to showing love and respect. An

TABLE I
Definitions of Types of Bias and Equivalence

Concept	*Definition*
Bias	Nuisance factors that threaten the comparability of scores
Construct bias	Construct measured is not identical across groups
Method bias	Nuisance factors result from factors such as sample incomparability (sample bias), instrument characteristics (instrument bias), and tester effects and communication problems (administration bias)
Item bias	Nuisance factors are at the item level
Equivalence	Comparability of test scores across cultures
Construct inequivalence	Comparisons lack a shared attribute ("comparing apples and oranges")
Structural or functional equivalence	Instrument measures the same construct in all of the groups studied
Measurement unit equivalence[a]	Measurement scales have the same units of measurement but different origins (e.g., the Celsius and Kelvin scales in temperature measurement)
Scalar (or full score) equivalence[a]	Scores are fully comparable across cultures (same origin and measurement unit in all cultures); this does not imply that means are identical across cultures

[a]Assumes interval- or ratio-level measurement in all cultures studied.

inventory of filial piety based on the Chinese conceptualization will cover aspects unrelated to the concept among Western individuals, whereas a Western-based inventory will leave important Chinese aspects uncovered. Construct bias can also be caused by differential appropriateness of the behaviors associated with the construct in the different cultures. Depression is often taken to consist of somatic aspects (e.g., loss of appetite, sleeplessness) and psychological aspects (e.g., loss of motivation, pessimism). Cross-cultural studies show that the psychological symptoms usually show up as a single cluster, whereas the somatic symptoms do not always form a single cluster or do not form a cluster that is related to the somatic symptoms.

2.2. Method Bias

Method bias is a label for all sources of bias emanating from the method and procedure of a study, including factors such as sample incomparability, instrument differences, tester and interviewer effects, and mode of administration. There are three types of method bias: sample bias, administration bias, and instrument bias. The first type, sample bias, refers to confounding sample differences. Research on the cognitive differences between literate and illiterate individuals, for example, has been plagued by sample bias because a comparison of literates and illiterates is nearly always a comparison between schooled and unschooled persons. So, a study of the influence of literacy almost inevitably becomes a study of the influence of schooling. A few studies have been carried out in cultures where literacy in a local script is transmitted in other than school settings. It is interesting to observe that these studies (which successfully avoided sample bias) found smaller differences between literates and illiterates than did the older studies in which schooled and unschooled individuals were compared. Sample bias tends to increase with the cultural distance between the samples.

Administration bias, the second source of method bias, can be caused by differences in the procedures or mode used to administer an instrument. For example, when interviews are held in participants' homes, physical conditions (e.g., ambient noise, presence of others) are difficult to control. Participants and clients are more prepared to answer sensitive questions in (anonymous) self-completion questionnaires than in the "shared" discourse of an interview. Another source of administration bias is ambiguity in the questionnaire instructions and/or guidelines or a differential application of these instructions (e.g., answers to open questions are considered to be ambiguous and require follow-up questions). The effect of test administrator or interviewer presence on measurement outcomes has been studied empirically. Larger effects have been found in interview studies of attitudes than in standardized testing of cognitive abilities. Deference to the interviewer has been reported; participants were more likely to display positive attitudes to a particular cultural group when they were interviewed by someone from that group. A final source of administration bias is constituted by communication problems between the

TABLE II
Sources of Bias in Cross-Cultural Assessment

Type of bias	*Source of bias*
Construct bias	• Only partial overlap in the definitions of the construct across cultures
	• Differential appropriateness of the behaviors associated with the construct (e.g., skills do not belong to the repertoire of one of the cultural groups)
	• Poor sampling of all relevant behaviors (e.g., short instruments)
	• Incomplete coverage of all relevant aspects/facets of the construct (e.g., not all relevant domains are sampled)
Method bias	• Incomparability of samples (e.g., caused by differences in education or motivation)[a]
	• Differences in environmental administration conditions, either physical (e.g., recording devices) or social (e.g., class size)[b]
	• Ambiguous instructions for respondents and/or guidelines for administrators[b]
	• Differential expertise of administrators[b]
	• Tester/Interviewer/Observer effects (e.g., halo effects)[b]
	• Communication problems between respondent and tester/interviewer (including interpreter problems and taboo topics)[b]
	• Differential familiarity with stimulus material[c]
	• Differential familiarity with response procedures[c]
	• Differential response styles (e.g., social desirability, extremity scoring, acquiescence)[c]
Item bias	• Poor item translation and/or ambiguous items
	• Nuisance factors (e.g., item may invoke additional traits or abilities)
	• Cultural specifics (e.g., incidental differences in connotative meaning and/or appropriateness of the item content)

Source. Adapted from van de Vijver and Tanzer (1997).
[a]Sample bias.
[b]Administration bias.
[c]Instrument bias.

tester and the participant or client. Language problems may be a potent source of bias when, as is not uncommon in cross-cultural studies, a test or an interview is administered in the second or third language of one of the parties involved.

The third source of method bias is instrument bias, which involves general features of an instrument that lead to unintended cross-cultural differences. Individuals from various cultures do not deal in the same way with Likert-type rating scales such as 5- or 7-point scales for assessing agreement. For example, Hispanics have been found to show more extreme scores than have European Americans on a 5-point scale but not on a 10-point scale.

In addition to the construct and the method, items can be the source of bias. Item bias, also known as differential item functioning, is a generic term for all disturbances at the item level. According to a definition that is widely used in psychometrics, an item is biased if respondents who have the same standing on the underlying construct (e.g., equally intelligent), but who come from different cultures, do not have the same mean score on the item. For instance, if a geography test administered to pupils in Poland and Japan contains the item "What is the capital of Poland?," Polish pupils can be expected to show higher scores on this item than can Japanese students, even when pupils with the same total test score are compared. The item is biased because it favors one cultural group across all test score levels. Several psychometric techniques are available to identify item bias. The most common sources of item bias are poor item translation, ambiguities in the original item, low familiarity or low appropriateness of the item content in certain cultures, and the influence of cultural specifics such as nuisance factors or connotations associated with the item wording. For example, in a translation of the American word "aggression," it is difficult or even impossible to maintain the combined meaning of violence (e.g., "an aggressive predator") and enterprising energy (e.g., "an aggressive salesperson") of the original.

3. EQUIVALENCE

Equivalence refers to the comparability of test scores obtained in various cultural groups. It involves the question of whether scores obtained in different cultures can be compared in a meaningful way. Equivalence and bias are related. If scores are unbiased (free from nuisance factors), they are equivalent and can be compared across cultures.

Four different types of equivalence can be distinguished: construct inequivalence, structural equivalence, measurement unit equivalence, and scalar equivalence. Construct inequivalence amounts to "comparing apples and oranges." Many examples of inequivalent constructs can be found in the clinical literature on culture-bound syndromes. For example, "Latah," a syndrome that can be found only among some Asian groups, consists of a sudden fright, resulting in uncontrollable imitative behaviors (e.g., verbal repetition of obscenities).

An instrument administered in various cultural groups shows structural equivalence if it measures the same construct in all of these groups. Statistical techniques, notably factor analysis, are employed to examine structural equivalence. If an instrument yields the same factors in various cultural groups, there is strong evidence that the instrument measures the same underlying construct(s). Structural equivalence does not presuppose the use of identical instruments across cultures. For example, a depression measure may be based on partly or entirely different indicators in each cultural group and still show structural equivalence. Structural equivalence has been addressed for various cognitive tests and personality measures.

The third type of equivalence is measurement unit equivalence. Instruments show this type of equivalence if their scales have the same units of measurement but different origins (e.g., the Celsius and Kelvin scales in temperature measurement). This type of equivalence assumes interval- or ratio-level scores (with the same measurement units in each culture). It applies when a source of bias with a fairly uniform influence on the items of an instrument affects test scores of various cultural groups in a differential way. Social desirability and stimulus familiarity may exert this influence. If these factors have a differential influence on the scores obtained in various cultural groups, observed score differences are hard to interpret due to the confounding of valid score differences and measurement artifacts. Empirical research shows that the role of response sets (in personality and attitude measurement) and stimulus familiarity (in cognitive testing) cannot be neglected. For example, there are country differences in social desirability; more specifically, because social desirability is inversely related to national wealth and educational level, individuals who come from richer countries or who are better educated tend to show lower scores on social desirability.

Only in the case of scalar (or full score) equivalence can direct comparisons be made using statistical tests such as the *t* test and analysis of variance (ANOVA). This is the only type of equivalence that allows for the conclusion that average scores obtained in two cultures are different or equal. Scalar equivalence assumes that identical interval or ratio scales are applicable across cultural groups.

Structural, measurement unit, and scalar equivalence are hierarchically ordered. The third type presupposes the second type, which presupposes the first type. As a consequence, higher levels of equivalence are more difficult to establish. It is easier to verify that an instrument measures the same construct in various cultural groups (structural equivalence) than to identify numerical comparability across cultures (scalar equivalence).

4. CULTURAL FACTORS IN THE ASSESSMENT OF PERSONALITY AND INTELLIGENCE

There are various examples of emic models of personality. These do not provide a pan-cultural view of personality but rather try to describe personality in a specific cultural context as accurately as possible. For example, Sow developed a model of African personality in which personality consists of various layers. Important elements in the model are spirituality and the role of ancestors and the extended family. Doi described a Japanese model of personality in which "amae" plays an important role. Amae refers to "indulgent dependency" and the passive love and dependence that find their origin in the love of mother and child and that form the template of relationships later in life. Because these studies are not comparative, bias and equivalence might not seem to be an issue. However, from a slightly more distant perspective, these studies address the bias in Western models because they indirectly show where Western models of personality fall short. It is suggestive that in both models (as well as in other non-Western models of personality), social and relational aspects of personality play a more central role than is common in Western models of personality.

A few Western-based personality inventories have been administered in many countries. A notable one is the Eysenck Personality Inventory. According to Eysenck, personality consists of three aspects: neuroticism, psychoticism, and extraversion. The inventory, used to assess these factors, shows these three factors with remarkable consistency. The factors underlying the instrument tend to be stable, indicating that the structural equivalence of the instrument is good. Another currently popular personality instrument is based on the Five-Factor Model of personality. Extensive research has shown that the five factors in this model that are assumed to underlie personality (Extraversion, Neuroticism, Agreeableness, Conscientiousness, and Openness) are remarkably consistent across cultures. However, studies of the measurement unit equivalence of the scales tend to produce more ambiguous results. Statistical tests (e.g., confirmatory factor analysis) do not provide clear support for measurement unit equivalence, but it is unclear to what extent this problem is due to the high sensitivity of the statistical tests employed to test the equivalence or to real differences in the measurement unit across cultures.

It can be concluded that cross-cultural research on personality has shown that the Western models can claim a broader generality than their cultural contexts of origin. Structural equivalence is strongly supported. It should be noted that studies have been restricted to schooled populations, so it is still unclear to what extent these results also apply to illiterates. It is still unclear to what extent non-Western theories and models could add other universal factors. Recent studies suggest that the Five-Factor Model of personality may be incomplete. Researchers from Hong Kong found a sixth factor, Interpersonal Relatedness, in China and replicated its existence in a Fijian sample. If new factors emerge and it turns out that Western models are incomplete, it is likely that the new factors will be located in the social and relational domains.

The domain with probably the most widely documented cross-cultural research is intelligence. Standard (Western) intelligence tests have been administered in a wide variety of cultures. In many multicultural societies, these tests have been administered to majority group members and usually at least one of the larger ethnic minority groups. Instruments such as Raven's Progressive Matrices and the Wechsler intelligence scales for adults and children have been found to reveal a remarkably stable factor structure across cultures, supporting the structural equivalence of intelligence as measured by these tests.

Statistical tests of measurement unit and full score equivalence of intelligence test scores tend to produce more ambiguous results. It is often difficult to decide whether equivalence in a given case is scalar equivalence or measurement equivalence. For example, the debate about ethnic differences in cognitive test performance can be largely seen as focusing on the level of equivalence of cross-cultural score comparisons and on the presence of biasing factors (notably method bias). For example, Herrnstein and Murray argued that when appropriate intelligence tests are used, cross-cultural differences in test performance reflect valid intergroup differences and show full scalar equivalence. The implication is that group differences in the intelligence test scores reflect differences in intellectual abilities, whereas group differences reflect mainly (or exclusively) method bias in the reasoning of their critics. Many of these critics argue that common intelligence tests show problems such as differential familiarity and that this method bias will allow only measurement unit equivalence.

Despite the disagreement about the higher levels of equivalence of cognitive test scores, cross-cultural studies show that it is very unlikely that intellectual processes are fundamentally dissimilar in various cultures. The findings are not supportive of theories that hold that different cultures produce their own kinds of intelligence; instead, they are more supportive of theories according to which cultures can shape the same set of human cognitive processes in a differential way. However, recent cross-cultural work points to another possible one-sidedness of Western tests. These tests tend to focus on two aspects of intelligence: reasoning and memory. Implicit theories of intelligence have been investigated in a number of countries by asking participants which characteristics they associate with intelligence. Social aspects are nearly always found to be important, yet social intelligence is not part of all intelligence tests. It is interesting to note that the social domain has also been mentioned in relation to personality research as an area where expansion of Western models might be needed to accommodate non-Western models.

Thus, cultural factors affect mental measurement in various ways. Some of these factors, such as a poor item translation, may be artifactual. Other factors, however, are more basic and may refer to problems with the whole instrument or even with the concept measured by the instrument. It is clear that methodological and substantive considerations cannot be separated when dealing with cultural factors in mental measurement. Knowledge of both mental measurement and the culture

being dealt with is essential in establishing quality in cross-cultural mental measurement.

See Also the Following Articles

Attitude Measurement ■ Conformity across Cultures ■ Cross-Cultural Psychology, Overview ■ Intelligence Assessment ■ Measurement and Counseling ■ Personality Assessment

Further Reading

Barrett, P. T., Petrides, K. V., Eysenck, S. B. G., & Eysenck, H. J. (1998). The Eysenck Personality Questionnaire: An examination of the factor similarity of P, E, N, and L across 34 countries. *Personality and Individual Differences*, *25*, 805–819.

Berry, J. W., Poortinga, Y. H., Segall, M. H., & Dasen, P. R. (2002). *Cross-cultural psychology: Research and applications* (2nd ed.). Cambridge, UK: Cambridge University Press.

Bond, M. H. (Ed.). (1996). *Handbook of Chinese psychology*. Hong Kong: Oxford University Press.

Camilli, G., & Shepard, L. N. (1994). *Methods for identifying biased test items*. Thousand Oaks, CA: Sage.

Cheung, F. M., Leung, K., Zhang, J., Sun, H., Gan, Y., Song, W., & Xie, D. (2001). Indigenous Chinese personality constructs: Is the Five-Factor Model complete? *Journal of Cross-Cultural Psychology*, *32*, 407–433.

Church, A. T. (2001). Personality measurement in cross-cultural perspective. *Journal of Personality*, *69*, 979–1006.

Church, A. T., & Lonner, W. J. (1998). The cross-cultural perspective in the study of personality: Rationale and current research. *Journal of Cross-Cultural Psychology*, *29*, 32–62.

Cole, M. (1996). *Cultural psychology: A once and future discipline*. Cambridge, MA: Harvard University Press.

Dana, R. H. (Ed.). (2000). *Handbook of cross-cultural and multicultural personality assessment*. Mahwah, NJ: Lawrence Erlbaum.

Herrnstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. New York: Free Press.

Jahoda, G. (1982). *Psychology and anthropology: A psychological perspective*. New York: Academic Press.

Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.

Johnson, T. P., & Van de Vijver, F. J. R. (2003). Social desirability in cross-cultural research. In J. A. Harkness, F. J. R. Van de Vijver, & P. P. Mohler (Eds.), *Cross-cultural survey methods* (pp. 195–204). New York: John Wiley.

Matsumoto, D. (Ed.). (2001). *Handbook of culture and psychology*. Oxford, UK: Oxford University Press.

Prezworski, A., & Teune, H. (1970). *The logic of comparative social inquiry*. New York: John Wiley.

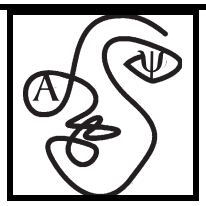
Sinha, D. (1997). Indigenizing psychology. In J. W. Berry, Y. H. Poortinga, & J. Pandey (Eds.), *Handbook of cross-cultural psychology*, Vol. 1: *Theory and method* (2nd ed., pp. 129–169). Boston: Allyn & Bacon.

Suzuki, L. A., Ponterotto, J. G., & Meller, P. J. (Eds.). (2001). *Handbook of multicultural assessment: Clinical, psychological, and educational applications* (2nd ed.). San Francisco: Jossey–Bass.

Van de Vijver, F. J. R., & Leung, K. (1997). *Methods and data analysis for cross-cultural research*. Thousand Oaks, CA: Sage.

Van de Vijver, F. J. R., & Tanzer, N. K. (1997). Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology*, *47*, 263–280.

Yamaguchi, S. (2001). Culture and control orientations. In D. Matsumoto (Ed.), *Handbook of culture and psychology* (pp. 223–243). Oxford, UK: Oxford University Press.



Mentoring

Joyce E. A. Russell
University of Maryland, College Park, Maryland, USA

1. Definition of Mentoring
2. Functions Provided by Mentors
3. Phases in Mentoring Relationships
4. Outcomes of Mentoring
5. Challenges with Mentoring Relationships for the Organization
6. Suggestions for Improving Mentoring Relationships in the Organization
7. New Directions for Mentoring Research
8. Conclusion
Further Reading

GLOSSARY

acceptance and confirmation The feeling that both parties (mentor and protégé) have toward each other whereby they feel comfortable in expressing their views.
challenging tasks The demanding or stimulating work assigned by the mentor, who gives feedback to help the protégé grow on the job.
coaching The career advice provided by the mentor to the protégé; the mentor suggests strategies and tactics for accomplishing work objectives and reaching career goals.
cultivation When career and psychosocial functions are enhanced during the 2 to 5 years after the mentoring is initiated.
initiation When the mentoring relationship begins to develop during the first year; both the mentor and protégé see it as mentoring.
redefinition The phase after the separation phase when the mentoring relationship may end or change qualities (e.g., the protégé becomes more like a peer).
role modeling The protégé's interest in modeling himself or herself after the mentor.
separation When the mentoring relationship changes after the cultivation period; typically, the protégé becomes more independent and is less needing of the mentor's career assistance.
sponsorship When the mentor nominates the protégé for promotions and/or important lateral moves either verbally or simply by association with the mentor.

Mentoring refers to a developmental relationship that exists between two persons, typically a more experienced senior person and a less experienced junior person. The mentoring relationship can also exist among peers or others within and outside an organization. Most mentoring relationships develop according to four phases: initiation, cultivation, separation, and redefinition. Mentors serve two primary functions: fostering career development of their protégés and providing psychosocial support. Numerous benefits have been reported from mentoring relationships for both mentors and protégés as well as for the organization.

1. DEFINITION OF MENTORING

Traditionally, mentoring has been defined as a developmental relationship that occurs between a more experienced senior employee and a less experienced junior employee. This involves a single dyadic relationship that often exists within one organization. However,

©2004 Elsevier Inc.
All rights reserved.

given the changing nature of organizations today (e.g., use of more project teams, flatter organizational structures), mentoring relationships are also formed among peers within and outside the organization. Mentoring generally consists of a close relationship between two people, with the primary goal of helping one of them (i.e., the protégé) develop in his or her career. This enables the protégé to address career challenges and take advantage of opportunities to grow on the job. In addition, the relationship is often a mutually enhancing one given that the mentor also receives some benefits (e.g., protégé support, work done by the protégé).

2. FUNCTIONS PROVIDED BY MENTORS

It is commonly accepted that mentors serve two basic types of functions to protégés: career development and psychosocial support. These are described in the following subsections.

2.1. Career Development

Career development refers to those aspects of the relationship that enhance the protégé's career advancement or growth. The mentor is able to help the protégé due to the mentor's experience, organizational rank, and/or influence in the organization. As a result of these functions, the protégé is often able to learn the "ropes" of the organization, gain exposure, and obtain promotions more quickly. In fostering career development, the mentor may serve any of the following roles:

- *Sponsorship.* The mentor nominates the protégé for promotions and/or important lateral moves either explicitly or simply by association with the mentor. The mentor opens doors and makes connections for the protégé that will advance his or her career.
- *Exposure and visibility.* The mentor gives the protégé opportunities to interact with senior managers or other important people through either written proposals or oral presentations. By initiating relationships with other key people, the protégé becomes visible to them and gains access to future opportunities.
- *Coaching.* The mentor provides career advice to the protégé and suggests strategies and tactics for accomplishing work objectives and reaching career goals. The mentor might provide feedback on assignments and offers ideas to improve work. In addition, the mentor helps the protégé to know who the key players are in the organization and describes strategies for working with them successfully. Essentially, the mentor teaches the protégé the ropes for how to be successful in the organization.
- *Protection.* The mentor guards the protégé from potentially damaging conflicts or problems with others. The goal is to protect the protégé's reputation from harm. Thus, the mentor might take the blame for mistakes that are beyond the protégé's control or might act as a buffer. The mentor might limit the protégé's exposure to senior people if the mentor believes that the protégé is not experienced enough to meet with them.
- *Challenging tasks.* This usually refers to mentoring provided by a manager to a subordinate. The manager assigns demanding or stimulating work and gives feedback to help the protégé grow on the job. The mentor could also place the protégé on various project teams to help expand the latter's knowledge and skills.

2.2. Psychosocial Support

Psychosocial support refers to those aspects of the relationship that enhance the protégé's competence, identity, self-worth, and effectiveness in a professional role. The close interpersonal relationship that exists between the mentor and the protégé encourages the development of trust and intimacy and enables the various roles to be fulfilled. These include the following:

- *Counseling.* This refers to the mentor serving as a sounding board to the protégé so that the latter can share with someone who will listen to his or her concerns and doubts about self, work, and family. The mentor provides empathetic listening and serves as a confidant or as someone who provides advice and encouragement.
- *Acceptance and confirmation.* This refers to the feeling that both parties (mentor and protégé) have toward each other that allows them to feel comfortable in expressing their views. The feeling of acceptance by the mentor enables the protégé to feel that he or she can try new things and speak candidly. The mentor shares positive feedback that he or she has heard about the protégé. The mentor sends encouraging messages (e.g., e-mail, phone) to the protégé to help him or her deal with pending challenges.
- *Friendship.* The mentor and protégé spend time together in a nonwork setting or in leisure activities or more relaxed settings (e.g., meals, sports, outings).
- *Role modeling.* This refers to the protégé's interest in modeling himself or herself after the mentor. In this case, the protégé has great respect and admiration for the

mentor and identifies with him or her. The protégé tries to emulate the mentor's relationships with others or the ways in which the mentor works on tasks. The mentor explains what he or she is doing and why. This is so common among mentors and protégés that some researchers consider it a third (separate) function in addition to career development and psychosocial support.

3. PHASES IN MENTORING RELATIONSHIPS

Mentoring relationships typically develop according to four phases: initiation, cultivation, separation, and redefinition.

1. *Initiation.* During the first year, the relationship begins to develop and the mentor and protégé both see it as becoming a mentoring relationship. The mentor provides career development (e.g., coaching, visibility) and some psychosocial support (e.g., role modeling). Typically, both individuals sense a feeling of excitement and challenge as they begin to work together. The protégé feels that someone is assisting him or her, and the mentor feels that he or she is taking on someone to groom, develop, and coach.

2. *Cultivation.* During the next 2 to 5 years, the career and psychosocial functions are enhanced. The mentor and protégé both benefit from the relationship and develop a stronger connection. However, not all mentoring pairs will develop in the exact same way. It depends on the needs and interests of the mentor and protégé. In general, both parties view this period in a positive light.

3. *Separation.* In general, after the cultivation period, the relationship changes. Typically, the protégé becomes more independent and is less dependent on the mentor for career assistance. If the protégé has achieved the same organizational level as the mentor, this limits the opportunities that the mentor can provide. In addition, if the mentor experiences job obsolescence or plateauing, this limits the type of career support that he or she can provide to the protégé. Similarly, if the individuals are no longer working at the same geographical location or in similar industries, they might drift apart. During the separation phase, there may be some feelings of sadness or confusion as the relationship ends or changes.

4. *Redefinition.* After the separation phase, the relationship may end or change qualities. For example, the protégé and mentor may become more like peers, and they may redefine their relationship as being more of a friendship. There may be some ambivalence and discomfort as they alter their relationship. At the same time, the protégé may feel a great sense of self-confidence and autonomy by "making it" to the next level.

4. OUTCOMES OF MENTORING

Research has indicated numerous positive benefits associated with mentoring for the protégé, the mentor, and the organization. These are described in the following subsections.

4.1. Benefits for the Protégé

The mentoring relationship enhances the protégé's personal and professional development, including subjective and objective indicators of career success. Through enhanced career opportunities (e.g., by sponsorship, by exposure and visibility), the protégé acquires critical learning experience that helps him or her to be able to demonstrate competence on the job. The protégé also gains an appreciation for the informal networks and political processes in the organization. This has typically led to greater career progress, faster promotion rates, and greater career satisfaction. It has been well documented that, when comparing individuals who had mentors with those who did not, protégés report significantly less work–family conflict, greater job satisfaction, greater salary growth, lower turnover intentions, and greater perceived career success. They also have greater socialization into the organization by gaining valuable information about the firm's practices.

4.2. Benefits for the Mentor

The mentor also reaps benefits from the mentoring relationship. By sponsoring the protégé for important positions or promotions in the organization, the mentor also gains. If the protégé is successful, others view the mentor as having made the right choice, and this enhances the mentor's credibility in terms of finding and developing talent in the organization. In addition, if the mentor's views are adopted by the protégé, this ensures that the mentor's perspectives will be carried forward in the organization for the future. By providing challenging work to the protégé and delegating to him or her, the mentor gains valuable technical assistance and support that enables him or her to be able to work

on other projects. In addition, the mentor has the loyalty of the protégé and is often kept informed by him or her about new work developments that facilitate the mentor's job performance. The psychosocial support that the mentor provides to the protégé also makes the mentor feel important and valued. If the protégé looks up to the mentor and tries to emulate him or her, this makes the mentor feel respected and appreciated in the workplace. If the mentor is asked to serve as a confidant or "counselor," this may make the mentor feel helpful and enhance his or her feelings of accomplishment. The mentor's satisfaction may also increase due to helping a less experienced colleague.

4.3. Benefits for the Organization

The organization also benefits from the mentoring relationship. In fact, of the rated "100 best companies to work for" in the United States each year, the majority indicate that they have established formal mentoring programs. One benefit to the organization is that protégés are better socialized into the corporate culture because mentors spend time teaching them the ropes. This means that protégés learn about the organization's values and performance metrics more quickly. In addition, mentoring is also related to lower protégé turnover intentions. This higher retention can save the firm considerable money and time that would have been spent recruiting new employees.

5. CHALLENGES WITH MENTORING RELATIONSHIPS FOR THE ORGANIZATION

There are a number of challenges associated with having a formal or informal mentoring program in the organization. Some of these are described in the following subsections.

5.1. Role Modeling Identification

As noted earlier, one important component of mentoring relationships is the degree to which the protégé identifies with the mentor. In some cases, this may be relatively easy due to similar backgrounds or life experiences. In other cases, it may be more problematic. For example, research has indicated some potential problems with the mentoring relationship where the mentor and protégé differ in their gender or race. In both cases, there may be difficulties with the protégé being able to identify with the mentor and vice versa. For example, a female protégé who has children and is trying to balance work and family may have some trouble relating to an older male mentor who has a wife who takes care of all of the family issues. Likewise, the female protégé's experiences in being given stereotypical tasks may be difficult for her male mentor to understand. Thus, a female protégé may get more support from a female mentor who has experienced similar life and career issues. Interestingly, male protégés often are also better able to identify with role models of the same gender. Consequently, the outcomes that protégés may reap from mentoring relationships may differ depending on whether they are in cross-gender or same-gender mentoring relationships. Similarly, an African American protégé who is placed on numerous committees at work to make sure that there is "diversity of views" may have difficulty in expressing his or her concerns to a White mentor who has never had to deal with tokenism issues. Generational differences in values and concerns may also make it more difficult for protégés to relate to mentors in the organization. For example, "Generation X" protégés who are more concerned with leisure and family pursuits may have more difficulty in identifying with the serious work ethic of their "baby-boomer" mentors.

5.2. Perceptions by Others of the Mentoring Relationship

In some cases, there are no problems with the mentoring relationship, yet others' perceptions may cause problems. For example, with a cross-gender relationship, others may make false assumptions about the nature of the relationship (i.e., assuming it is personal/sexual). This might force the mentor and protégé to curb their social activities (e.g., having lunch, playing golf) and, thus, limit the extent to which psychosocial support or friendships can be nurtured. Consequently, they may limit their interactions to a professional or work setting only. In addition, a cross-gender relationship might cause rumors to fester about why the protégé was given various opportunities. Peer resentment and public scrutiny might force the protégé and mentor to deal with new issues and dilemmas.

5.3. Establishing Formal Versus Informal Relationships

Formal mentoring relationships are those established by the organization. Typically, mentors and protégés

are matched up based on some criteria such as similarity of interests and experiences. In contrast, informal mentoring relationships evolve spontaneously and without a formal mandate. Either the mentor or the protégé seeks the other one out or they both initiate the relationship. These informal relationships are considered to be voluntary, whereas formal mentoring is considered to be prescribed by the organization. In general, research has indicated that informal mentoring results in more benefits (e.g., greater support), possibly due to the closer bonds that are formed in the relationships. Yet having formal assigned relationships ensures that everyone receives some type of mentoring. Thus, the challenge for the organization is to design effective formal relationships while still facilitating the development of informal relationships.

5.4. Risks Associated with Mentoring

There may be risks associated with mentoring someone in the organization. For example, a mentor who sponsors or provides exposure and visibility to a protégé who performs poorly in front of senior managers might make the mentor look bad (i.e., the mentor's judgment in picking this person may be called into question). In addition, a mentor who is overly protective of a protégé by holding him or her back from potential opportunities can limit the career advancement of the protégé. Research has indicated that this is more likely to happen in relationships where the mentor and protégé are of a different gender or race. A senior White male mentor might be worried about showing off a female or African American protégé because he assumes that the protégé will be watched very closely and that if he or she does not do well, it will reflect poorly on the mentor. This may keep the protégé from developing critical learning experiences to grow on the job. When establishing mentoring relationships in the organization, it is important to train mentors and protégés on what their roles should be in these relationships and on what obstacles they may encounter and how to deal with them.

5.5. Dysfunctional Mentoring Relationships

Often, the assumption is that all mentoring relationships are positive or of high quality during all phases of development. Recent research has shown that this is not the case. Sometimes, mentoring relationships can be dysfunctional or only marginally effective. For example, cases have been reported where protégés receive assistance that actually derails their careers. In other cases, there are instances of deception, sabotage, harassment, bullying, jealousy, gamesmanship, and/or poor performance by either mentors or protégés. More research is needed to further define the types of behaviors that are considered dysfunctional and their implications for protégés, mentors, and the organization. In addition, strategies to address these problems and enhance the quality of the mentoring relationship need to be defined.

6. SUGGESTIONS FOR IMPROVING MENTORING RELATIONSHIPS IN THE ORGANIZATION

6.1. Make Sure That Everyone Has Access to Mentoring

Research indicates that not everyone has equal access to mentoring, yet given its importance for the success of protégés, the organization should ensure that everyone can have mentors. Creating opportunities for social activities and cross-functional work teams can facilitate the development of informal mentoring relationships. In addition, the formation of formal mentoring relationships can ensure that everyone is assigned to a mentor. The use of mentoring circles to bring multiple senior people together with several junior or mid-level people can enable a diverse group of individuals to work together.

6.2. Encourage Individuals to Establish Multiple Mentoring Relationships

Protégés should form relationships with several mentors and not rely solely on one mentor to fulfill all career and psychosocial functions. Sometimes, aligning with only one mentor can prove to be troublesome if protégés are required to follow only that one person's ideas and advice. This gives protégés only one perspective on how to be successful. It would be more advantageous if they formed alliances with several mentors at various points in time. It has been shown that professionals with multiple sources of mentors have higher levels of career success and feelings of career satisfaction than do those with only a single mentor or with no mentor at all. In fact, research on successful Black

executives indicates that they draw on multiple sources for mentoring rather than rely on a single mentor. It is also important for mentors to form relationships with several protégés. This enables them to gain support and technical assistance from multiple sources and also reduces charges of favoritism.

6.3. Establish a Culture That Encourages Mentoring Relationships

Mentoring requires time to develop productive relationships. An organization that rewards individuals for developing each other (based on the performance management and incentive systems) will enable mentoring to flourish. However, a firm that focuses only on "bottom line" results and has no incentives for employees to spend time assisting others will discourage mentors from taking on protégés. Likewise, establishing formal social opportunities (e.g., lunches, celebrations) and work activities (e.g., teams) will enable the formation and growth of informal developmental relationships, some of which could lead to mentoring relationships. The organization should also support individuals' efforts to establish relationships outside of the firm. This is particularly important given the trend today for employees to work at a variety of organizations rather than just one firm during their lifetimes.

7. NEW DIRECTIONS FOR MENTORING RESEARCH

There is considerable research in the field of mentoring today. Some of the current topics for investigation were described previously, and other ideas for future research are listed in the following subsections.

7.1. Clarify the Mentoring Construct

Traditionally, mentoring referred to a single dyadic relationship between a senior person and a junior person. This made mentoring relatively easy to study because it was clearly defined in this manner. Recently, many new ideas have been generated about what mentoring refers to within the broader array of developmental relationships and networks that people form. As such, the degree to which mentoring differs from coaching, support groups, advising, and counseling needs to be better explicated.

7.2. Examine Alternative Forms of Mentoring

Although it has always been suggested that mentoring can take a variety of forms, during the past few years, researchers have encouraged more investigation into these different types of relationships. Future work should examine the nature of peer relationships as well as other developmental networks. It seems clear that people use a variety of relationships to gain both career and psychosocial assistance, yet more research is needed to understand how they develop these relationships and what benefit these relationships have for protégés and mentors over time.

7.3. Explore Mentoring for Diverse Types of Employees

Today's rapid changes in technology and organizational structures, as well as the emergence and decline of companies, open up new challenges for workers trying to get ahead or stay marketable. Research is needed to study mentoring among some of these workers.

7.3.1. Cross-Generations

Given differences in the values and work ethic among employees of various generations, it seems important to view what cross-generational mentors and protégés are looking for in these relationships. Do mentor–protégé pairs from different generations have unique obstacles and difficulties? If so, how can these be addressed?

7.3.2. Free Agent Workers

These refer to independent employees whose goal is to be marketable and flexible enough to move where the opportunities present themselves. Research is needed to determine the type of mentoring they need and seek. Compared with traditional workers, are they more inclined to seek mentoring from a variety of sources and networks?

7.3.3. Entrepreneurs

With the greater numbers of new business start-ups, especially among women, it will be increasingly more important to examine the types of assistance and mentoring that are needed to help them be successful.

7.3.4. Global Workforce

As firms continue to "go global," it is important to examine the type of career counseling, mentoring, and training that employees need to be successful in international assignments. For example, research has indicated that mentoring for women managers in international assignments is critical for their career success. What type of mentoring is needed, and how does it differ depending on the global assignment?

8. CONCLUSION

Mentoring involves a developmental relationship between two people that has benefits for the mentor, the protégé, and the organization. The career and psychosocial support that are given to the protégé enhance the personal and professional development of all parties involved. Although generally positive in nature, these relationships occasionally encounter difficulties and/or obstacles. With effective support and training provided from the organization, some of these issues can be better managed. Future research is needed to examine the underlying dynamics and processes involved in the mentoring relationship.

See Also the Following Articles

Career Counseling ■ Counseling Interview ■ Developmental Counseling

Further Reading

Arthur, M. B., & Rousseau, D. M. (Eds.). (1996). *The boundaryless career: A new employment principle for a new organizational era.* New York: Oxford University Press.

Feldman, D. C. (Ed.). (2002). *Work careers: A developmental perspective.* San Francisco: Jossey–Bass.

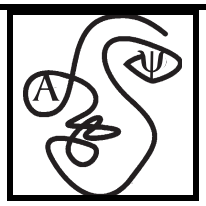
Hall, D. T. (2002). *Careers in and out of organizations*. Thousand Oaks, CA: Sage.

Higgins, M. C., & Kram, K. E. (2001). Reconceptualizing mentoring at work: A developmental network perspective. *Academy of Management Review*, *26*(2), 264–288.

Kram, K. E. (1985). *Mentoring at work: Developmental relationships in organizational life*. Glenview, IL: Scott, Foresman.

Ragins, B. R. (1997). Diversified mentoring relationships in organizations: A power perspective. *Academy of Management Review*, *22*, 482–521.

Russell, J. E. A., & Adams, D. M. (1997). The changing nature of mentoring in organizations: An introduction to the special issue on mentoring in organizations. *Journal of Vocational Behavior*, *51*(1), 1–14.



Models for Transportation

John D. Lee
University of Iowa, Iowa City, Iowa, USA

1. Introduction
2. Model Purpose
3. Model Scope
4. Model Type
5. Emerging Models of Driver Performance
6. Driver Performance Model in the Interactive Highway Safety Design Model
7. In-Vehicle Information Systems Demand Model
8. Rear-End Collision Warning Assessment Model
9. Cognitive Architectures for Modeling Driver Performance
10. Conclusions
Further Reading

GLOSSARY

cognitive architecture Provides a broad theory of cognitive capabilities and limits; when operationalized in Adaptive Control of Thought (ACT-R) or Soar, it provides a high-level programming language for simulating human cognition that includes many of the cognitive limits and constraints that guide human performance.

concurrent real-time model A model that generates output in parallel with driver behavior; this type of model may make it possible to infer driver state and intention when it is not possible to measure these variables directly.

driving performance The description of the driver and vehicle trajectory relative to important performance limits such as lane boundaries and speed limits; typical driving performance metrics include lane deviation, steering wheel reversals, speed variation, and brake reaction time.

driving safety The description of the likelihood of an event that compromises driver safety; typical driving safety metrics include crashes per million miles driven, injuries per million miles driven, and fatalities per million miles driven.

process model A model that predicts not only the outcome but also the response process that generates the outcome; this type of model is based on theoretical principles that go beyond the data used to develop the model.

product model A model that produces an output for given inputs but does not describe the causal links between the inputs to the model and the output; this type of model is typically based on interpolation of experimental data using linear regression, lookup tables, and neural nets.

Driver performance models can help to develop safe in-vehicle information systems by helping to coordinate experiments, evaluate systems, and modulate information flow. This article describes the dimensions of model purpose, scope, and transparency and then uses them to identify how several specific models may help to develop in-vehicle technology. As in-vehicle technology becomes increasingly complex and adaptive, driver performance models may need to address the effect of driver attitudes and preferences on behavior. This brief overview of four modeling approaches demonstrates some of the trade-offs in developing models of driver performance.

1. INTRODUCTION

This article focuses on models of driver performance that might address the challenges of integrating technology into the driving environment. Models of driver

©2004 Elsevier Inc.
All rights reserved.

performance can support evaluation of in-vehicle technology in several ways. First, models make it possible to examine design options before the systems are actually implemented. In this way, design concepts can be explored without the expense of developing prototype systems. Second, models can support extrapolation beyond experimental data. Experiments typically identify the relative merit of two to eight design options, whereas designers might need to understand the merit of hundreds of design options. Models can be used to estimate the merit of the design options that are not included in the experiments. Also, although catastrophic failures such as crashes are very rare and are unlikely to occur in an experiment, it may be important to understand how design parameters affect the frequency of those events. Models can help to extrapolate from measures of driving performance to measures of driving safety. For example, a model might help to translate measures of performance (e.g., lane deviations) to a measure of safety (e.g., accident rates per million miles driven). Beyond evaluating systems, models of driver performance can also play a useful role in the in-vehicle information systems themselves. Real-time models of driver capabilities can be used to modulate the information flow to the driver. In this way, a driver performance model can help to manage the flow of information so that it is consistent with the perceptual and cognitive limits of the driver.

These benefits of driver performance models are balanced by the challenges of developing such models. Fortunately, substantial research has focused on developing models of driver performance. For some applications, these models can be adapted and used without the cost of developing new models. Unfortunately, there is a great diversity of driver performance models, and understanding whether a particular model meets the needs of a particular evaluation is often unclear. This article describes the dimensions of model purpose, scope, and type and then uses these dimensions to describe how several specific models might be modified to address the demands of evaluating in-vehicle technology.

2. MODEL PURPOSE

Every model is created for a purpose. The purpose for which a model is developed has a major influence on the range of applications for which the model can generate useful results. The wide range of purposes of driver performance models demonstrates the challenge of finding a model capable of meeting the needs of a particular application. Driver models have been developed to investigate vehicle dynamics, suspension, and handling characteristics. They have also been developed to support the design of roadway geometry. Recently, models have also been developed to address driver response to in-vehicle technology such as warning systems and the visual and cognitive demands on in-vehicle information systems. Another class of models has been developed to address the safety implications of personality variables and their influence on driving behavior. For example, researchers have created statistical models that link risk-taking tendencies to driving safety. Others have created statistical models that predict accident involvement based on cognitive and perceptual deficits such as the driver's effective field of view.

3. MODEL SCOPE

Driving is a complex multifaceted task that has been modeled in many ways. For example, traffic models have been developed to describe the combined behaviors of many drivers. These models of traffic behavior play a critical role in understanding the effectiveness of in-vehicle technology aimed at traffic management and congestion reduction. Some models of traffic behavior describe driver behavior only at the aggregate level. These models often represent macro-level behavior by using principles of fluid dynamics to describe traffic flow. Other models of traffic behavior rely on micro-level description of the decision-making and response patterns of individual drivers that make up the traffic.

Beyond models of traffic flow, many models have been developed to describe the behavior of individual drivers. These models range from those that describe the entire range of driving activities to those that describe a small subset of activities (e.g., steering behavior for lane keeping) or a small subset of driving scenarios (e.g., response to rear-end collision situations). Michon defined three levels of driver activity: vehicle control (e.g., speed control, lane position), tactical maneuvering (e.g., speed choice, turn choice), and strategic decisions (e.g., choice of route, decision to initiate a trip). These three levels define qualitatively different types of behavior that might be included in a model. Models that address strategic decisions, such as route choice and trip timing, are often represented as statistical choice models. These model results provide important information for both marketing and design of in-vehicle information systems. Most models do not

combine a representation of strategic behavior with a representation of vehicle control or tactical maneuvering.

4. MODEL TYPE

In addition to distinguishing among models based on the scope of activities that they address, models can be distinguished based on whether they describe the driver's response process or just the outcome of that process. Models that describe only the outcome or product associated with a driving technology tend to be static models that pair an outcome with a set of inputs. More sophisticated static product models can combine many variables to predict driver performance in a wide range of situations; however, product models do not describe the causal links between the inputs to the model and the output. In contrast, a process model not only predicts the outcome but also specifies the response process that generates the outcome. In other words, process models are based on theoretical principles, whereas product models are based on interpolation, lookup tables, and curve fitting. Process models tend to be dynamic in that they take time-varying inputs and generate time-varying outputs. One benefit of process models is that they can predict driver responses over a range of conditions not included in the data used to create the models. Another benefit is that process models tend to make assumptions explicit. Rather than drawing on data that embody hidden assumptions, process models make assumptions easy to see because they are explicitly coded into the algorithms. Although dynamic process models have many advantages, they are not always feasible. In many cases, theory may be lacking, making it impossible to construct a process model. In addition, process models are often more complex, and the underlying mathematics may be less understandable. In contrast, product models have the advantage of providing simple solutions to complex problems. The cost of these simple solutions is the uncertainty regarding their accuracy and underlying assumptions.

5. EMERGING MODELS OF DRIVER PERFORMANCE

Several recent modeling efforts demonstrate the potential of driver models to address important questions regarding driver performance and vehicle design. The following discussion of these models does not present a complete survey of driver models but instead represents a sample of the more promising developments. In addition, these models illustrate the distinctions of model purpose, scope, and type that are important for developing a modeling strategy.

Table I summarizes the three models, showing three distinctly different approaches to modeling driver performance. The first model, the Driver Performance Model (DPM), is an ambitious attempt to model the dynamics of driver behavior. This process model produces a time history of vehicle position and speed as well as a time history of driver state. The second model, Design Evaluation and Model of Attention Demand (DEMAnD), is a comprehensive product model that identifies a level of demand associated with a task performed with an in-vehicle information system such as entering a destination for a route navigation system using a touch screen. The third model, the Attention-based Rear-end Collision Avoidance Model (ARCAM), is a much less ambitious process model that generates a time history of a driver's response to the very specific scenario of emergency braking to avoid an imminent rear-end collision. The following discussion also briefly addresses the potential of general cognitive architectures to act as a foundation for a driver performance model.

6. DRIVER PERFORMANCE MODEL IN THE INTERACTIVE HIGHWAY SAFETY DESIGN MODEL

DPM will help designers to understand the demands that a particular curve or other aspect of roadway geometry places on the driver. DPM is a very ambitious model that intends to capture both vehicle control (e.g., speed and path control) and tactical maneuvering aspects of driving (e.g., speed and path choice). This broad scope is bounded by a focus on driver response of road geometry of two-lane rural roads. DPM is a detailed process model. It generates predictions of driver behavior based on algorithms that are tied to elements of human cognition that are relevant to driving. These algorithms are based on a theoretical framework that relates cognitive resources, such as attention, perception, and motor control, to the driving demands associated with moving down a virtual roadway. As a process model, DPM produces a response to state

TABLE 1
Characteristics of Three Representative Driver Models

	Purpose	*Scope*	*Type*
DPM	Support geometric design of roadways	The control and tactical aspects associated with speed and path maintenance on a two-lane rural road	Process model based on a nested closed-loop feedback control
DEMAnD	Support interface design of in-vehicle information systems	The control aspects of interacting with in-vehicle technology Considers driving only implicitly as it is degraded by competition for common attention resources	Product model based on lookup tables and linear regression
ARCAM	Support selection of appropriate algorithm parameters for a rear-end collision warning system	The control and tactical aspects of driver response to imminent collision situations	Process model based on a field theory of driving and a combination of open-loop and closed-loop control
Cognitive architecture	Model high-level cognition in a variety of settings	Aspects of cognition that correspond to control and tactical aspects of driving	Process model based on empirical data collection and experimental validation

changes that are updated at each time step of the simulation. This results in time history of the driver state and vehicle position on the road. Because the developers created DPM using a general framework with a strong theoretical basis, it is feasible to extend DPM to go well beyond its current purpose. Adapting it to address distraction associated with in-vehicle devices seems feasible.

7. IN-VEHICLE INFORMATION SYSTEMS DEMAND MODEL

The emergence of a wide array of in-vehicle technology has promised increased safety, mobility, convenience, and efficiency. The purpose of DEMAnD is to help engineers evaluate design options before creating a prototype in-vehicle information system. More specifically, this system enables engineers to compare two competing options, evaluate a modification to an existing system, and evaluate a task or design feature against safety criteria.

The focus of DEMAnD is on in-vehicle tasks, and driving performance is modeled indirectly; that is, driving performance decrements are estimated based on the demands of tasks associated with in-vehicle information systems. DEMAnD takes the tasks and subtasks associated with an interaction with an in-vehicle device and produces estimates of the attentional demands of those tasks. DEMAnD summarizes the consequence of these attentional demands in terms of expected decrements in driving performance. Unlike DPM, DEMAnD does not generate a time history of driver actions and driving performance; instead it combines data across in-vehicle tasks to estimate a general indicator of the safety consequence. Because the model does not generate output for the entire time history of the interaction among the driver, the system, and the roadway, it might not reflect how peaks in attentional demand could overwhelm the driver.

DEMAnD is a static product model of driver performance. It uses lookup tables and linear regression equations to relate particular design alternatives to declines in driving safety. Unlike DPM, DEMAnD does not generate a time history of driver response to events; instead it adds the estimated cognitive demands of a set of tasks and subtasks. This approach does not attempt to describe the dynamic process that the driver goes through in interpreting and responding to an in-vehicle information system.

8. REAR-END COLLISION WARNING ASSESSMENT MODEL

ARCAM was developed to aid designers in evaluating candidate collision warning strategies. This model predicts driver response to warnings produced by various collision warning algorithms in different driving contexts (e.g., initial velocity, following distance, deceleration of the lead vehicle, initial velocity of the lead vehicle). With these predictions, designers can compare various algorithms of parameter choices for a wide range of roadway conditions. Currently, ARCAM predicts behavior of drivers only in imminent rear-end collision situations.

Figure 1 shows the major components of ARCAM. A central aspect of ARCAM is the focus of driver attention. A driver's attention to the roadway depends on the uncertainty about the state of the roadway. This uncertainty increases until it reaches an unacceptable level, at which time the driver returns attention to the road. The uncertainty grows as a function of information density of the road, vehicle velocity, rate of forgetting, and time. The following equation describes the driver's uncertainty:

$$\text{Standard Deviation of Memory Position} = \alpha + \lambda \times t^{1.5}, \qquad (1)$$

where t is the time since the last sample, α is the initial uncertainty associated with the driver's estimate of collision potential, and λ is the rate of uncertainty growth.

ARCAM is a detailed process model. It generates predictions of driver behavior based on algorithms that are tied to elements of human cognition that are relevant to drivers' response to imminent collision situations. These algorithms are based on a field theory of perception and a model of attention allocation that seems to capture the most critical cognitive processes associated with collision avoidance. As a process model, ARCAM produces a response to state changes that is updated at each time step of the simulation. This results in a time history of the driver state and vehicle position on the

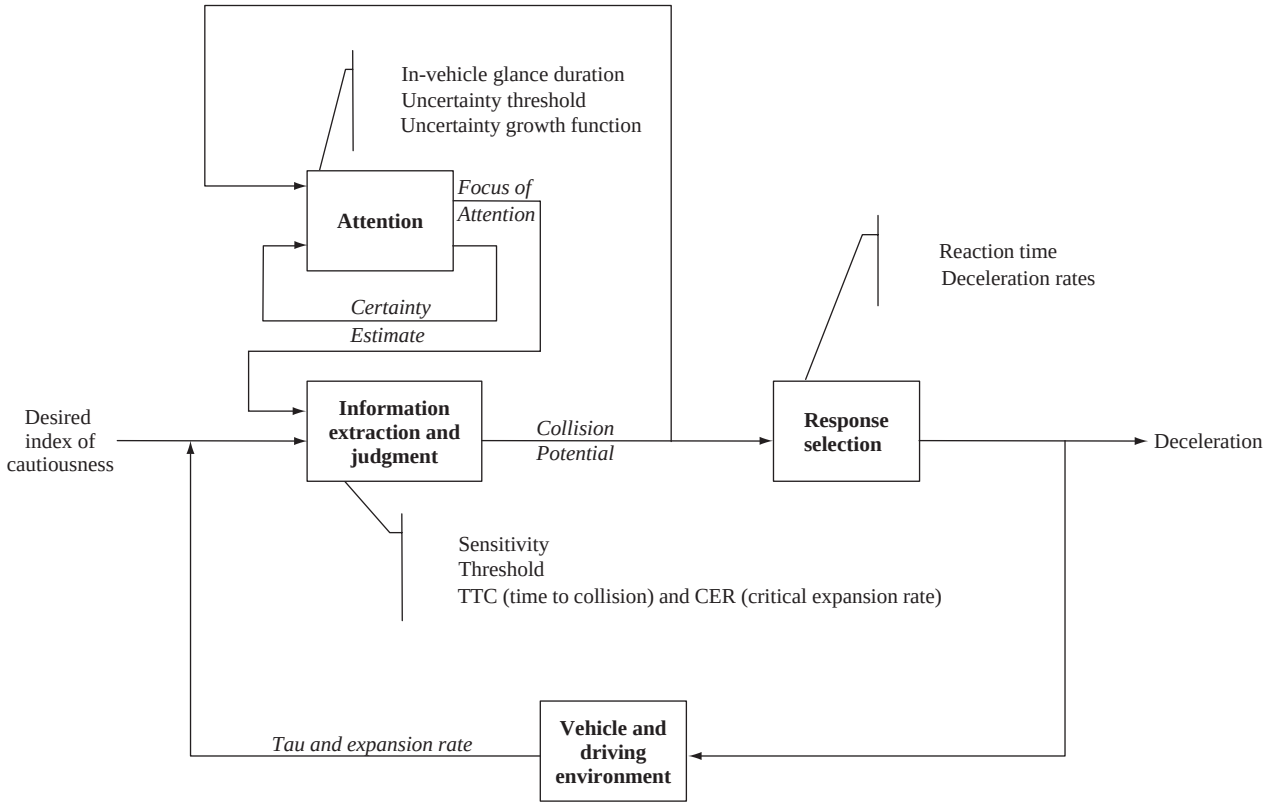


FIGURE 1 Major components of ARCAM. Reprinted from Brown and colleagues (2000).

road. This trajectory provides a basis for understanding how cognitive processes influence driver performance.

9. COGNITIVE ARCHITECTURES FOR MODELING DRIVER PERFORMANCE

In addition to models developed specifically to capture the demands of driving, several modeling tools or cognitive architectures could be used to model driving performance. These cognitive architectures include Adaptive Control of Thought (ACT-R), Soar, and Human Operator Simulation (HOS). A major benefit of these approaches is that they embody many of the cognitive limits and constraints of people. For example, ACT-R was developed to account for many elements of high-level cognition. ACT-R captures many aspects of how people learn and respond to problems. Using ACT-R or another cognitive architecture as a starting place for developing a driver model allows model developers to leverage many years of research as well as a validated approach to describing human behavior. Developing a model with ACT-R, Soar, or HOS can be much faster and provide results with greater validity than can an approach that develops the model without the benefit of an underlying cognitive architecture.

Although cognitive architectures offer important benefits, they also pose two substantial challenges. First, the level of detail assumed by many architectures might be more specific and detailed than many analyses need. Second, the cognitive architecture may make fundamental assumptions regarding human behavior that might conflict with how people drive. For example, ACT-R would describe driving in terms of discrete activities of symbol manipulation. Although this might be appropriate for some driving activities, many aspects of driving might best be described as continuous closed-loop control. Third, the tasks that many cognitive architectures have been designed to model are not necessarily those that comprise driving. Cognitive architectures act as a high-level programming language that can greatly enhance the efficiency of model development because they embody many of the cognitive constraints that govern driver performance. However, they have been developed to explain human performance in very different domains (e.g., human–computer interaction) and, thus, may make fundamental assumptions that conflict with the human performance capabilities and time-sharing strategies used for driving.

10. CONCLUSIONS

This brief overview of four modeling approaches demonstrates some of the trade-offs in developing models of driver performance. Comprehensive process models of driver performance, such as DPM, can be complicated to operate and resource intensive to create. However, they provide a powerful way in which to understand how design options affect driver cognition and performance. Outcome models, such as DEMAnD, can link large amounts of empirical data, using lookup tables and regression equations, to design parameters. They provide a powerful tool for designers, but they make many assumptions that might not be easily understood, and they might not be usable beyond the original purposes of the models. Focused process models, such as ARCAM, provide the benefits of the more comprehensive process models and are less resource intensive to build. Unfortunately, they might have such limited scope that they are useful only for only a very specialized application such as the evaluation of a rear-end collision warning algorithm. Cognitive architectures promise the generality of a comprehensive process model, such as DPM, without the demands of creating a complete modeling framework. The limitation of cognitive architectures such as ACT-R is that they have generally been developed for applications unrelated to driving, and the ability to adapt them to the driving domain has not been tested.

See Also the Following Articles

Accidents in Transportation ■ Aviation ■ Driving Safety ■ Traffic Safety Assessment ■ Transportation Systems, Overview ■ Travel Behavior and the Environment

Further Reading

Anderson, J. R. (1990). *Adaptive character of thought*. Hillsdale, NJ: Lawrence Erlbaum.

Ball, K., Owsley, C., Sloane, M. E., Roenker, D. L., & Bruni, J. R. (1993). Visual–attention problems as a predictor of vehicle crashes in older drivers. *Investigative Ophthalmology & Visual Science*, *34*, 3110–3123.

Brown, T. L., & Lee, J. D. (2000, January). *Driver modeling in the design of rear-end collision warning systems*. Paper presented at the annual meeting of the Transportation Research Board, Washington, DC.

Brown, T. L., Lee, J. D., & McGehee, D. V. (2000). An attention-based model of driver performance in rear-end collision situations. *Transportation Research Record*, *1724*, 14–20.

Byrne, M. D., & Anderson, J. R. (1998). Perception and action. In J. R. Anderson, & C. Lebiere (Eds.), *Atomic components of thought* (pp. 143–168). Mahwah, NJ: Lawrence Erlbaum.

Cameron, G. D. B., & Duncan, G. I. D. (1996). PARAMICS: Parallel microscopic simulation of road traffic. *Journal of Supercomputing*, *10*(1), 25–53.

Chowdhury, D., Santen, L., & Schadschneider, A. (1999). Vehicular traffic: A system of interacting particles driven far from equilibrium. *Current Science*, *77*, 411–419.

Goldbach, M., Eidmann, A., & Kittel, A. (2000). Simulation of multilane freeway traffic with detailed rules deduced from microscopic driving behavior. *Physical Review E*, *61*(2), 1239–1246.

Hankey, J. M., Dingus, T. A., Hanowski, R. J., Wierwille, W. W., & Andrews, C. (2000). *In-vehicle information systems behavior model and decision support: Final report* (FHWA-RD-00-135). McLean, VA: Turner–Fairbank Highway Research Center.

Jonah, B. A., & Dawson, N. E. (1987). Youth and risk: Age differences in risky driving, risk perception, and risk utility. *Alcohol, Drugs, and Driving*, *3*(3–4), 13–29.

Kantowitz, B. H., Levison, W. H., Hughes, W., Taori, S., Palmer, J., Dingus, T. A., Hanowski, R., Lee, J. D., Mears, B., & Williges, R. (1997). *Development of prototype driver models for highway design: Task A final Project I work plan.* Seattle, WA: Battelle Human Factors Transportation Center.

Klar, A., & Wegener, R. (1997). Enskog-like kinetic models for vehicular traffic. *Journal of Statistical Physics*, *87*(1–2), 91–114.

Klar, A., & Wegener, R. (1999). A hierarchy of models for multilane vehicular traffic I: Modeling. *Siam Journal on Applied Mathematics*, *59*, 983–1001.

Lane, N. E., Strieb, M. I., Glenn, F. A., & Wherry, R. J. (1981). The Human Operator Simulator: An overview. In J. Moraal, & K. F. Krais (Eds.), *Manned systems design methods, equipment, and applications* (pp. 121–151). New York: Plenum.

Levison, W. H. (1994). A simulation model for driver's use of in-vehicle information systems. *Transportation Research Record*, *1403*, 7–13.

Mannering, F. L., Abu-Eisheh, S. A., & Arnadottir, A. T. (1990). Dynamic traffic equilibrium with discrete/continuous econometric models. *Transportation Science*, *24*, 105–116.

Mannering, F., Kim, S. G., Ng, L., & Barfield, W. (1995). Travelers' preferences for in-vehicle information systems: An exploratory analysis. *Transportation Research C: Emerging Technologies*, *3*, 339–351.

Michon, J. A. (1989). Explanatory pitfalls and rule-based driver models. *Accident Analysis and Prevention*, *21*, 341–353.

Modjtahedzadeh, A., & Hess, R. A. (1993). A model of driver steering control behavior for use in assessing vehicle handling qualities. *Journal of Dynamic Systems Measurement and Control*, *115*, 456–464.

Moray, N. (1981). The role of attention in the detection of errors and the diagnosis of errors in man–machine systems. In J. Rasmussen, & W. B. Rouse (Eds.), *Human detection and diagnosis of system failures*. New York: Plenum.

Morrall, J., Miller, E., Smith, G. A., Feuerstein, J., & Yazdan, F. (1995). Planning and design of passing lanes using simulation model. *Journal of Transportation Engineering*, *121*(1), 50–62.

Nelson, P., & Raney, B. (1999). Objectives and benchmarks for kinetic theories of vehicular traffic. *Transportation Science*, *33*, 298–314.

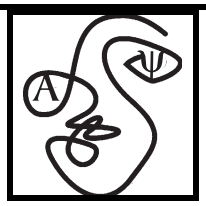
Newell, A. (1992). Precis of unified theories of cognition. *Behavioral and Brain Sciences*, *15*, 425–492.

Reagan, J. A. (1994, Summer). The Interactive Highway Safety Design Model: Designing for safety by analyzing road geometrics. *Public Roads*, pp. 37–43.

Schuller, J., Schaefer, F., Neukum, A., & Krueger, H. P. (1999). A driver model as development tool for vehicle handling design. *Vehicle System Dynamics*, *33*, 110–121.

Senders, J. W., Kristofferson, A. B., Levison, W. H., Dietrich, C. W., & Ward, J. L. (1967). The attentional demand of automobile driving. *Highway Research Record*, *195*, 15–33.

Sharp, R. S., Casanova, D., & Symonds, P. (2000). A mathematical model for driver steering control, with design, tuning, and performance results. *Vehicle System Dynamics*, *33*, 289–326.



Motivational Taxonomies

Jeffrey B. Vancouver
Ohio University, Athens, Ohio, USA

1. Introduction
2. Definitions
3. Content Theories
4. Process Theories
5. Implications for Work Settings
Further Reading

GLOSSARY

contingency Pairing or relationship between two events, behaviors, or thoughts; usually, the pairing or relationship is probabilistic (i.e., it does not always happen).

intrinsic motivation Behavior performed for its own sake and not for material or social rewards.

justice Perceptions of the fairness of the allocation of outcomes (i.e., distributive justice) or policies and procedures used to make the allocation decisions (i.e., procedural justice).

Motivating oneself or others in a work setting requires understanding many aspects of human thought and behavior. Several theories have emerged to facilitate understanding what people want and how they learn and process information to get what they want.

1. INTRODUCTION

Motivation has long been a topic of interest to applied psychologists. In work settings, motivational processes are thought to relate to decisions of participation, such as joining or leaving an organization, and to performing well in an organization. Motivation is also implied in decisions and actions related to broader organizational concerns such as unionization, constructive behaviors (e.g., speaking well of one's employer), and destructive behaviors (e.g., workplace violence). An understanding of motivational constructs and processes can help employers to predict who will be motivated to join and work hard, why some are not so motivated, and what can be done to increase participation and performance generally and for specific individuals. From an employee's perspective, an understanding of motivation might help the individual to make participation decisions that lead to satisfying work experiences and techniques for motivating oneself to perform desired tasks promptly and well. To address these issues, the field of work motivation has generated a list of theories that speak to specific issues. This article describes some historically important and currently prominent theories in the field.

2. DEFINITIONS

Before reviewing the theories, it is useful to define the two concepts comprising the title of this article: motivation and taxonomy. Motivation is difficult to define. It is easier to understand in terms of what it is not: ability, knowledge, and skills. For example, a well-written essay requires knowledge of the topic, the ability to write, and the willingness to write. Only the willingness element is

©2004 Elsevier Inc.
All rights reserved.

motivation. More directly, motivation is often defined as that which determines the form, direction, intensity, and duration of one's actions. However, one must recognize that ability or other influences (e.g., situational constraints) might at least partially determine these results. Thus, separating motivational causes from nonmotivational ones represents a continual dilemma for motivational researchers.

Taxonomy represents a classification scheme in which elements are assigned to no more than one class and all of the elements are assigned; that is, it is mutually exclusive and exhaustive. Taxonomies are also often hierarchical; that is, within a class, subclasses are often identified. Here, content may be used as one higher order class and process as another class. Content theories deal with a particular property or set of properties thought to influence motivation. They address the question of what people want or need. Process theories focus on how individuals handle information or content. They address the question of how people learn to obtain or process information to obtain what they want and need. Process theories are content free; that is, they usually assume that people are hedonistic (i.e., seek pleasure and avoid pain), but they make no statements about what people inherently find to be pleasurable and painful. Content theories usually have some description of processes, but they also include descriptions (often taxonomies) of what might give people pleasure and pain.

3. CONTENT THEORIES

3.1. Needs Theories

One well-known taxonomy of human desires (or needs) was presented by Maslow, a clinical psychologist, based on a lifetime of interaction with patients. This taxonomy was arranged in a hierarchy designed to represent the order of precedence (Fig. 1). The bottom level (physiological needs) describes what individuals need to survive. Higher level needs describe what individuals need to thrive. A process element to Maslow's theory was that lower level needs had to be met to some extent before higher level needs became motivating. Maslow's hierarchy was important for work motivation because it included in the higher order needs more than just the material needs. Indeed, the highest need (self-actualization) describes the desire of individuals to be all that they can be. It represents a desire to strive, in both personal and work lives, to achieve new heights, to create, and to produce to the best of one's ability. This sentiment resonated with managers and executives because they believed that these kinds of desires motivated their behavior.

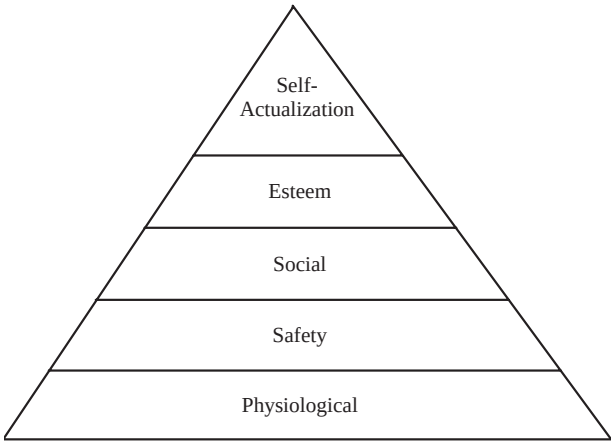


FIGURE 1 Maslow's hierarchy of needs.

Unfortunately for Maslow's theory, the majority of empirical research did not support his hierarchy. Some suggested that the hierarchy be reduced to two or three levels with some processes added. Yet even these changes were not reconcilable with empirical and logical counterarguments (e.g., an individual can be eating a sandwich to meet a physiological need while writing his or her self-actualizing masterpiece). These issues led most motivational researchers to abandon the specific framework but to scavenge useful pieces from the hierarchy, including the general idea that people desire more than what money can buy.

3.2. Job Design

One area of motivational research that benefited from the needs approach was job design. If applied psychologists could understand what was intrinsically motivating about jobs, they might be able to redesign them so that workers were more motivated and more satisfied. Interviews with workers seemed to reveal that although lack of material rewards and/or poor working conditions (i.e., extrinsic elements) could make them unhappy, it was the opportunity to do a meaningful job well that was truly motivating and satisfying.

So, what makes a job meaningful? The most prominent theory in this area, called the job characteristics model, claims that it is the degree to which the job involves a variety of skills used to create a whole important thing (or service). In addition, the job should allow the individual (or team) to decide what, when, and how

to do the thing (or provide the service). Finally, the job should provide feedback regarding progress directly to the job holder. To the extent that these factors are present, an individual doing the job should be motivated because it allows the individual to demonstrate competence and an ability to affect his or her world. The theory also recognizes that if these needs in the individual are not strong, these job dimensions are not likely to be important and might even aggravate the individual.

Because job design theories are specific about what should be done, they are easier to test than are vague needs taxonomies. In this case, the empirical literature tends to be much more supportive of the recommendations within the theory, particularly in terms of job satisfaction. In terms of helping an organization's bottom line (i.e., profit and productivity), the findings are less clear. One issue is that recommendations from motivation-oriented job design techniques and engineering-oriented job design techniques often conflict. Engineering-oriented job designers tend to focus on efficiency, error reduction, and reliability. This approach emerged from Ford's assembly line formula that seemed to greatly improve the efficiency of production through work simplification, even as it substantially reduced the humanity of work. Some now suggest balancing all of the relevant criteria (e.g., reliability, intrinsic motivation) when designing and evaluating jobs.

3.3. Equity and Justice

A different class of theories focuses on what an individual does not want, that is, to be treated unjustly. Simply observing elementary school-aged children distributing a reward among themselves reveals the importance of this desire for fairness (and the difficulty in achieving it). The desire does not disappear with adulthood. Organizational members seem acutely aware of the fairness of distributions of outcomes and the policies used to make these distributions. The original theory focused solely on the distributions of outcomes (i.e., distributive justice). More recent work has focused on the fairness of distribution policies (i.e., procedural justice).

Actually, the original justice theory, called equity theory, was focused on a process, that is, the comparison between the outcomes that an individual receives and the outcomes that others receive. Moreover, these outcomes were appraised in relation to the inputs used to derive them. Hence, the theory predicted that if an individual received $5 for a job and a friend received $10, the former would perceive this as fair if the friend worked twice as long or twice as hard. On the other hand, if the individual perceived that both he or she and the friend put equal work into the job, that person would see the pay as unfair and be motivated to rectify the situation. The individual might do that by working less hard in the future, finding some other form of compensation (e.g., pilfering office supplies), denigrating the friend's work to the boss, or coming to believe that he or she does not really put in the same quality of work as the friend. If none of these produced a sense of justice, the person might quit that job. Research has found that all of these things can happen. Meanwhile, the friend is likely to believe that his or her efforts are somehow worth more than the first person's efforts, relieving any sense of overpayment inequity. Researchers hoped that individuals would be motivated to work harder to overcome their sense of overpayment inequity, but the evidence seems to indicate this is rarely the case.

More recent research has focused on the perceived policies and procedures for making distribution decisions. Voice in the process, whether at the policy development stage, the policy implementation stage, or a follow-up stage (e.g., ability to appeal decisions), seems to be a central issue. In addition, thoughtful explanations for negative outcomes (e.g., layoffs) seem to reduce justice-restoring behaviors (e.g., sabotage). The procedural justice research seems to be summed up in the concept that one should treat people with respect.

4. PROCESS THEORIES

Process refers to how people meet (or strive for) their wants and needs. Three kinds of process theories have gained prominence in organizational psychology: learning theories, decision theories, and goal theories. Learning theories describe how people learn to achieve their desires. Decision theories describe how people process information such that they make choices leading them to meet their desires, including how they choose goals. Goal theories describe the properties of goals that most influence behavior and performance.

4.1. Learning Theories

Most learning theories claim that people make associations between responses and the consequences that follow those responses. If the consequences are what people want, people are more likely to make those responses in the future. Thus, if making a big sale earns the respect of bosses and colleagues, an individual is more likely to engage in the activities that led to the big sale in the

future. In addition, when stimuli (e.g., events) are frequently paired, people learn to associate the stimuli, allowing them to predict their environment to some extent. Thus, one might associate certain attributes of customers with sales. Observations of customers with such attributes indicate opportunities not otherwise present. The result might be the engagement of certain behaviors associated with sales directed at customers with certain attributes.

In organizational settings, the most common application of learning theories is behavior modification programs based on reinforcement theory. They involve determining the desired behaviors of employees and then setting up contingencies with rewards when the behaviors are observed. Well-designed studies of these programs have generally supported them, but several aspects of the approach make it onerous for organizations. First, it requires that one or a few specific behaviors are what all employees should be doing. For many jobs, this might not be the case. Second, it requires someone (or something) to look for the behaviors by the employees. Such extensive monitoring can be expensive and aversive to employees. Third, because the theory is content free (i.e., it does not state what people want), it requires figuring out what is rewarding for each employee. This generally is not a problem because most people want pay, which is a very flexible means to other ends. Unfortunately, employee pay is an organizational cost that organizations would like to minimize. Hence, some effort is made at finding less expensive rewards (e.g., praise). Fourth, theorists in this area advocate using the programs to encourage (i.e., reinforce) desired behaviors as opposed to discouraging unwanted behaviors. This can limit the approach's applicability. Finally, some motivation theorists argue that behavior modification programs undermine intrinsic motivation because they do not allow individuals to infer competence from their behavior.

In response, learning theorists point out that learning is constantly occurring in organizations. Organizations could simply hope that their work environments encourage the behaviors desired of employees, or they could attempt to construct environments that encourage the desired behaviors (i.e., behavior modification programs). Moreover, the intrinsic motivation issue does not seem to be a problem because individuals use the rewards to infer competence. Nonetheless, because of the difficulties associated with identifying an easily monitored set of specific behaviors, many researchers have sought to identify important properties of the work environment that might encourage or discourage the desired behaviors. For example, the beliefs and desires of coworkers have been found to be significant sources of influence on learning and behavior.

4.2. Decision Theories

Decision theories have several advantages over other theories of motivation from the perspective of motivational researchers. First, as the key dependent variable, they can use self-report measures of choice rather than behavior, which can be hopelessly confounded with ability. Second, because decision theories assume the cognitive processing of information, researchers can also obtain reports from workers about the contingencies and associations they perceive in their environments. These theories focus on what individuals do with the associations learned as opposed to how they are formed. Hence, researchers can be less concerned with what the environments look like and can focus on how those environments are interpreted by the workers.

Several decision theories exist, but within organizational psychology, the expectancy–value models have received the most attention. Expectancy–value models reduce workers' interpretations of their environments to three key constructs. First, they are concerned with the contingencies that workers see between specific behaviors and performance; that is, individuals' belief in the probability that behaviors will lead to performance (although these theories consider more than performance). Second, expectancy–value theories claim that the associations that individuals form between performance and other outcomes are important. Finally, expectancy–value theorists want to know what anticipated value workers place on the outcomes associated with performance. Given this information, the theories predict that individuals will choose to engage in behaviors that will lead to performance (or choices) if the outcomes associated with performance (or choices) are positively valued. The amount of effort or persistence directed toward the behaviors will weaken to the extent that individuals believe the behaviors might not achieve performance or believe the performance might lead to negative or nonvalued outcomes.

The predictions from decision theories and learning theories are very similar. They both advocate the setting up of contingencies between performance and valued rewards. However, they are more likely to encourage training or the removal of performance barriers if behavior-to-performance expectancies are low instead of setting up specific behavior-based contingencies. This moves the monitoring to the level of performance and away from specific behaviors.

Research on decision theories is generally supportive. People do seem to use their beliefs regarding contingencies and value to make choices regarding what jobs to take and other work-related behaviors. Moreover, self-report methods of eliciting these beliefs can be useful as organizations seek to understand why performance levels are not up to standard or what might improve them. However, for more mundane issues, such as determining how much effort to exert toward performance, the theories are less impressive. Indeed, exceptions seem to be the rule, and decision researchers are attempting to explain why people deviate from doing what seems to be in their best interest as defined by an expectancy–value theory.

4.3. Goal Theories

One decision that seems to be critical to understanding, predicting, and influencing performance is the choice of a performance goal. Most current research focuses on goal theories, which continue the cognitive orientation found in the decision theories. That is, goals are defined as desired states internally represented within the mind. Properties of this representation, including its level (e.g., produce a little or a lot), its level of specificity (e.g., do one's best or do 20), and one's commitment to it seem to substantially affect goal-striving behavior. Specifically, difficult (i.e., high-level), well-specified goals with strong commitment will lead to higher performance than will easy vague goals with low commitment. Although the research has been somewhat ambiguous regarding the commitment concept, the level and specificity constructs are well supported within boundary conditions. For example, feedback regarding performance progress is required for goals to have an effect. Moreover, learning and complex tasks might be adversely affected by difficult performance goals, but learning goals or contexts where the goals encourage the development of appropriate strategies for performing complex tasks can enhance these outcomes. Also, it seems clear that through the integration of decision theories, beliefs about contingencies with valued outcomes can enhance commitment and the acceptance of difficult goals.

4.4. Integrative Efforts

Although controversy surrounds the value or potential of integrating the various motivational theories, several efforts have been made in that direction. One early attempt combined expectancy–value notions with a particular content, that is, need for achievement. According to this theory, achieving difficult tasks (where expectancy was low) is more rewarding but may also lead to failure. Depending on each individual's desire for achievement, fear of failure, and expectancies, these factors predict choice and effort. Unfortunately, the evidence was not always supportive, and the theory has lost favor. However, the content (i.e., need for achievement) remains a candidate for being a measure of general work motivation.

More recently, social cognitive theory has emerged as an integrating theory of the process models of motivation. Social cognitive theory grew out of social learning theories that included concepts of learning from experience and from watching others as well as associations derived from people's history of experiences. Social cognitive theory, like expectancy–value theories, added the importance of how individuals perceive and interpret events and contingencies. For instance, individuals might use their feelings of physiological arousal to help them judge their ability to perform behaviors. According to the theory, a person's belief in his or her capacity to engage in a specific behavior, called self-efficacy, is key to understanding how the individual allocates resources to the task and how he or she reacts to frustrations and successes. For instance, individuals who are high in self-efficacy are likely to raise their goals when they achieve existing goals. Hence, goal processes are also included in social cognitive theory.

5. IMPLICATIONS FOR WORK SETTINGS

Several implications of specific theories have been described in this article. This section addresses some specific applied questions and what the theories do and do not have to say about them.

5.1. Who Should Be Hired?

One job of applied psychologists is to find individual differences that predict (i.e., correlate with) job performance. If measures of these predictors can be given during the application phase, organizations can use them to select individuals who will perform well. This strategy has proven to be fairly successful when it comes to measuring and predicting ability, knowledge, and skill. However, predicting who will be motivated has proven to be more difficult than predicting ability. This makes sense. Content theories (e.g., need for

achievement), which are most likely to provide explicit measures, have often suffered due to measurement problems. Even if the measures were reliable and valid, highly motivated individuals might direct their effort to nonwork endeavors (e.g., hobbies). Moreover, jobs often leave little leeway regarding what should be done or how to do it. For these reasons, research has tended to focus on more specific measures of motivation (e.g., work values) that indicate a fit with the organization where the job allows some autonomy or ambiguity.

5.2. Why Do Individuals Accept Job Offers and Stay?

Assuming that an organization likes what it sees in an applicant, the applicant has the next choice with regard to the organization (i.e., whether to join). Later, after being hired, the employee may decide to quit. Decision theories have much to say about these processes, and they have been used extensively to predict these outcomes. Particularly with regard to the decision to join an organization, factors such as pay and benefits play an important role. In regard to quitting, these factors and the working environment have an impact. Of particular importance for these decisions are perceptions of equity and fair treatment.

5.3. Why Do Individuals Work Hard (or Not)?

The motivational process theories argue for making pay and other rewards contingent on performance and providing specific difficult goals. Although the second recommendation is generally well received, the first one has created controversy. It seems to help motivate some employees, and some even like the system. However, in other situations, the measures of performance needed for contingency plans to work are considered suspect. If the measures (a) do not capture the behavior or performances that lead to organizational productivity or (b) are perceived to be a function of influences outside of an individual's control (e.g., the politics and idiosyncracies of supervisors), the measures are unlikely to work well and can lead to unwanted effects. Yet if individuals were informally surveyed about their performance appraisal systems, one would likely find few employees who believe that the systems capture all that they are supposed to capture and that they capture only what they are supposed to capture. Adding insult to injury is that in jobs where performance is difficult to measure, it is also difficult to set fair performance goals. For instance, consider a high school teacher. How should performance for this job be measured? What goals could be set that do not rely on input from the students that is beyond the teacher's control?

Indeed, can the possibilities expressed in Maslow's view of people be realized? Can organizations and jobs be designed to be places where workers can derive meaning and satisfaction? Can policies and procedures that are perceived as just be developed? Can a fair wage be provided for work done without having to monitor the work done? These are challenging questions for which the motivational theories provide some reasons for optimism as well as some reasons for pessimism. This leads to the final, and perhaps most important, question: Can a better understanding of human motivation help to answer these questions? Among work motivation researchers, the answer is uniformly "yes."

See Also the Following Articles

Intrinsic Motivation and Self-Determination ■ Job Analysis, Design, and Evaluation

Further Reading

Ambrose, M. L., & Kulik, C. T. (1999). Old friends, new faces: Motivation research in the 1990s. *Journal of Management*, *25*, 231–292.

Campbell, J. P., & Pritchard, R. D. (1976). Motivation theory in industrial and organizational psychology. In M. D. Dunnette (Ed.), *Handbook of industrial and organizational psychology* (pp. 63–130). Chicago: Rand McNally.

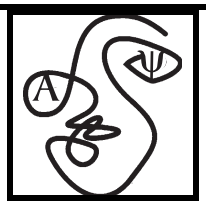
Donovan, J. J. (2002). Work motivation. In N. Anderson, D. Ones, K. Sinangil, & C. Viswesvaran (Eds.), *The handbook of industrial, work, and organizational psychology*. Thousand Oaks, CA: Sage.

Kanfer, R. (1990). Motivation theory in I/O psychology. In M. D. Dunnette, & L. M. Hough (Eds.), *The handbook of industrial/organizational psychology* (Vol. 1, pp. 75–170). Palo Alto, CA: Consulting Psychologists Press.

Landy, F. J., & Becker, W. S. (1987). Motivation theory reconsidered. *Research in Organizational Behavior*, *9*, 1–38.

Mitchell, T. R. (1997). Matching motivational strategies with organizational contexts. *Research in Organizational Behavior*, *19*, 57–149.

Pinder, C. C. (1998). *Work motivation in organizational behavior*. Upper Saddle River, NJ: Prentice Hall.



Motivation and Culture

Dennis M. McInerney
University of Western Sydney, Macarthur, Sydney, Australia

Martin L. Maehr
University of Michigan, Ann Arbor, Michigan, USA

Martin Dowson
Institute of Christian Tertiary Education, Round Corner, Australia

1. Cross-Cultural Approaches to Achievement and Achievement Motivation
2. Toward a Basis for Investigating Achievement and Achievement Motivation across Cultures
3. Applying Social Cognitive Conceptualizations of Motivation and Achievement
4. Methodological Improvements in Applied Settings
5. Conclusion
Further Reading

GLOSSARY

achievement motivation The desire to excel, that is, the impetus to strive for excellence and success.

attribution- and cognitive-style models Theoretical conceptions that emphasize the causal role of thinking patterns in determining motivation.

ecological research Research that suggests that the experiential and performance dimensions of events and situations are essential in understanding the achievements, behaviors, responses, and scores of groups and individuals.

emic research Research that develops culture-specific constructs and tools that may limit cross-cultural comparability.

etic research Research that develops universally appropriate constructs and tools and that enhances cross-cultural comparability but may limit access to emic data.

goal theory A theory that posits that the purposes students have for learning, such as doing better than others and improving their personal performance, influence the nature and quality of their motivation and engagement in learning.

personal investment theory A theory concerned with how persons choose to invest their energy, talent, and time in specific pursuits.

pseudo-etic research Research that assumes that research constructs and tools are etic when, in fact, they are emic; consequently, the validity of cross-cultural comparisons with pseudo-etic models is questionable.

social cognitive theory A theory that attributes learning to the cognitive processing of observed models of thought, action, and behavior.

The potential impact of culture on the nature and functions of psychological variables has long been recognized. However, until relatively recently, it was often assumed that definitions, methodologies, data, and generalizations from psychological research could be readily transferred, more or less, across cultural boundaries. Typically, this meant that psychological theory and research that was based in (and on) American or Western European settings was considered readily transferable to other cultures with few, if any, modifications. In particular, the assumption of transferability applied to a range of psychological constructs of interest in applied (i.e., real-life) settings. These constructs included intelligence, anxiety, expectancies, motivation, perceptions of success and achievement, and self-concept. However, there is now a wide recognition that such constructs are not necessarily structured in the same way across cultures. Moreover, the relative importance of various constructs may differ greatly among

©2004 Elsevier Inc.
All rights reserved.

cultures. For example, it has been demonstrated that the ways in which students construct their understandings of achievement and success (i.e., what "counts" for achievement/success and how achievement/success is gained) may be very different across cultures. For this reason, there has been considerable impetus to reformulate important psychological constructs (and research methodologies) to make them more relevant and appropriate to the specific cultures to be studied. This, in turn, suggests the need for development of taxonomies, nomenclatures, and conceptualizations of affect, cognition, and behavior that are culturally specific. This article begins by discussing early studies in, and conceptualizations of, achievement and achievement motivation. It then discusses more recent theoretical and methodological approaches to the study of these constructs, with particular emphasis being given to their application in cross-cultural settings.

1. CROSS-CULTURAL APPROACHES TO ACHIEVEMENT AND ACHIEVEMENT MOTIVATION

1.1. Early Studies in Motivation and Achievement

The study of achievement motivation owes much to the work of McClelland and associates. They suggested that achievement motivation was a personality trait that developed in some people more than in others as a result of early socialization experiences and the emotional concomitants of these experiences. Because socialization practices were presumed to vary across cultures, McClelland and colleagues' conceptualization provided a "ready-made" theoretical basis for the examination of achievement motivation in a range of groups, societies, and cultures. As a result, McClelland and colleagues' work gave impetus to many research programs investigating a range of issues related to achievement motivation in cross-cultural settings.

Specifically, early cross-cultural studies inspired by McClelland and colleagues typically examined the antecedent-consequent variables associated with the development of achievement motivation among diverse social and cultural groups. Some of these studies were particularly concerned with relationships between achievement motivation and variables such as racial/cultural identity, self-esteem, and academic achievement. Other studies were concerned with the sociocultural antecedents of achievement motivation, including independence training, social class, and acculturative stress.

1.2. Early Conceptualizations of Achievement and Achievement Motivation

McClelland and colleagues originally suggested that achievement behavior consisted of four distinct but related elements: competition with a standard of excellence, affective concern for goal attainment, an evaluation of performance, and some standard for the attainment of a long-term goal. These elements were typically identified in the projective responses to selected thematic apperception test (TAT) cards. In general, this approach categorized individuals as high or low in achievement motivation and determined those situations that would maximize their performance. Atkinson and colleagues introduced the concept of person (trait)/situation (state) interactions and explored the ways in which these interactions influence motivation and achievement. For example, Atkinson and colleagues' research suggested that high-achievement need individuals are likely to be motivated by achievement situations that are moderately challenging, that is, where success is perceived to be possible but requires effort. Alternatively, low-achievement need individuals are likely to be motivated in situations that are minimally challenging and, hence, require little perceived effort for success. Moreover, low-achievement need individuals may also be attracted to maximally challenging situations where perceived effort is great but where a reasonable "excuse" for failure (i.e., situation difficulty) is inherent. Atkinson and colleagues' research culminated in the formulation of expectancy–value theory, which systematically assessed the interaction between the value that individuals placed on perceived outcomes from achievement situations (value) and how likely they considered those outcomes to be (expectancies).

Overall, the research of McClelland and Atkinson and their colleagues suggested that achievement-oriented persons emphasized individual responsibility for their actions, were concerned with medium- to long-range goals, attributed success to their own abilities, were self-confident, and obtained most satisfaction from competitive situations where the outcomes were uncertain. Low achievement-oriented persons were characterized as being more motivated to avoid failure than to achieve success and as being most motivated by tasks that minimized the possibility of experiencing affect of failure.

This theme has been pursued up to, and including, studies by Bright in 1994 and Conchas in 2001 on student motivation and achievement.

Within its cultural context, the conceptualisation of McClelland and Atkinson and their colleagues was (and, to some extent, still is) quite an adequate interpretation of what constitutes achievement motivation. However, few studies questioned the basic premises of this conceptualization in terms of its cross-cultural validity. In contrast, most studies assumed that the conceptualization, its premises, and its associated methodologies were essentially etic (culturally transferable). Thus, when measures of differences between cultural groups on achievement and achievement motivation constructs were apparent, it was assumed that these reflected significant differences between groups on the constructs themselves rather than reflecting the applicability (e.g., appropriateness, salience) of the constructs across cultural groups. Interestingly, the non-Western cultural groups being examined often performed poorly with respect to given constructs. The conclusion drawn from such research was that "poor" performance reflected either (a) a deficiency in the group being examined (e.g., genetic inferiority, inadequate socialization for achievement) or (b) an incongruence and incompatibility between the values and norms held by the dominant cultural group and the assumptions, norms, values, and behaviors of the minority or culturally different group.

Perhaps more important, however, such studies often had the effect of "elevating" the goals, perceptions, and behaviors of Western cultural groups to the status of universal norms. In terms of cross-cultural work, this approach increasingly became recognized as both misguided and inadequate and has sometimes been termed "pseudo-etic." For this reason, more recent cross-cultural studies have explicitly recognized the potential impact of a variety of specific cultural components on achievement and achievement motivation. This culture-specific research is often referred to as emic research. As a result of this research, a significant amount of evidence has accumulated, indicating that modes of expression, achievement, and achievement motivation are not culturally invariant but rather very much related to, and embedded in, cultural contexts. In particular, the emphasis placed on individual achievement and success in McClelland and colleagues' theory is now questioned extensively. It is now argued that achievement must be analyzed in its total cultural context, including the various roles, both economic and noneconomic, through which achievement motivation may be directed. A wide range of studies in many different cultural settings support this view. A common theme identified in these studies is that socialization patterns identified as important to developing achievement motivation in Western contexts (e.g., independence, self-reliance) are clearly not manifest (or, at least, not manifest in the same ways or for the same reasons) in other cultural contexts. From this style of research, researchers are able to extract common features that appear to have universal validity and then compare cultures on the basis of these features. This approach is often referred to as etic research.

2. TOWARD A BASIS FOR INVESTIGATING ACHIEVEMENT AND ACHIEVEMENT MOTIVATION ACROSS CULTURES

Quite early in the development of cross-cultural research into achievement and achievement motivation, Kornadt and colleagues in 1980, in accord with Maehr in 1974, suggested that, in spite of the many potential configurations of achievement and achievement motivation, some general characteristics of achievement in all cultures may be identified, thereby emphasizing the etic dimensions of achievement motivation. They suggested that these characteristics might serve as a potential basis for cross-cultural work on achievement motivation. These components include the following:

- The existence of a standard of excellence for an individual's own goal-directed behavior
- The existence of affective reactions to success and failure (e.g., being proud and happy, being disappointed and sad)
- Individual feelings of responsibility for the outcome of achievement-related activity
- Personal incentives for achievement based on security or insecurity about an individual's own capacity or ability to succeed

Kornadt and colleagues suggested in 1980 that the culture-bound experiences, cognitive structures, and values of each individual will determine the way in which these components are manifest in particular cultures but that the components themselves are relatively robust across cultures. Thus, achievement-motivated behavior will appear in all cultures—but in different forms, stimulated by different cues, and directed at different goals.

The preceding framework acted (and still acts) as a useful starting point for investigating achievement and

achievement motivation across cultures. However, this framework needed to be supported by more detailed conceptualizations of achievement and achievement motivation. In response to this need, many theoretical models that attempted to define such conceptualizations more explicitly emerged. Some of these models continue to be particularly useful for research in applied cross-cultural settings. Four of these are outlined briefly in the following subsections: personal investment (PI) theory, ecological theory, attribution theory, and goal theory.

2.1. Personal Investment Theory

Maehr and Braskamp's PI model of achievement motivation builds on and integrates various dimensions from earlier conceptualizations of the nature of motivation that are still considered to be important in any explanation and analysis of motivation, particularly in cross-cultural settings. PI theory is concerned with how persons choose to invest their energy, talent, and time. PI theory is particularly relevant for investigations into how individuals of various cultural backgrounds relate to different achievement situations. This is because it does not assume that people from a given culture will invest effort in the same achievement situations or, if they do, for the same reasons as will those from other cultures. PI theory addresses the apparent paradox between the low assessment of individuals from non-Western cultural groups on achievement tests and their obvious achievement and motivation across a range of activities as illustrated by persistence, level of activity, high performance, and continuing motivation in chosen areas of behavior.

PI theory also emphasizes the role played by social and cultural contexts in determining motivational patterns in performing achievement tasks. Moreover, it is phenomenologically based and emphasizes the subjective meanings of situations in light of individuals' culturally determined belief systems such as beliefs about self, perceptions of appropriate goals, and perceived alternatives available for pursuing these goals.

PI theory is a social cognitive theory in that it assumes that the primary antecedents of choice, persistence and variations in activity levels, are thoughts, perceptions, and beliefs that are embedded in cultural and social beliefs about self and situation. Specifically, PI theory designates three basic components of meaning as critical to determining PI in specific situations:

- Beliefs about self, referring to the more or less organized collections of perceptions, beliefs, and feelings related to who one is
- Perceived goals of behavior in given situations, referring to the motivational focus of activity, particularly what the person defines as success and failure in these situations (e.g., task, ego, social solidarity, extrinsic rewards)
- Perceived alternatives for pursuing these goals, referring to the behavioral alternatives that a person perceives to be available and appropriate (in terms of sociocultural norms that exist for the individual) in a given situation

Each of these components of PI theory may be influenced differentially by the structure of tasks and situations, personal experience, and access to information, as well as, importantly, the sociocultural context in which tasks, situations, and persons are embedded.

As a model, PI theory alleviates many of the problems inherent in monocultural research models. In particular, it conceptualizes achievement motivation in terms that recognize the possibility of diverse modes of achievement behavior across cultures. PI theory also strikes a balance between personality and situation while incorporating dimensions (e.g., locus of control) that have been found to be useful in analyzing levels of achievement motivation. Studies by McInerney and colleagues in 1995 and 1997 illustrated the use of PI theory in applied cross-cultural settings.

2.2. Ecological Theory

At about the same time as PI theory was being developed, Berry in 1980 attempted to redress the balance in cross-cultural research between laboratory-based programs and research set in naturalistic environments by proposing an ecological model for cross-cultural studies. This model (or at least its general approach) continues to be relevant for contemporary cross-cultural studies. Berry suggested that research that is based largely on experimentation, with little consideration of the wide range of variables operating in natural environments, holds little value. To make the task of effective naturalistic research easier, Berry proposed a multilevel model. This model describes research that moves from a naturalistic–holistic level to a controlled reductionistic level through (a) four environmental contexts (ecological, experiential, performance, and experimental) and (b) four effects (achievements, behaviors, responses, and scores).

The ecological context (also referred to as the natural–cultural habitat) refers to all of the relatively permanent characteristics that provide a context for human action. The experiential context is the pattern of recurrent experiences that provide a basis for learning. The performance context is the limited set of environmental circumstances, immediate in space and time, that may be observed to account for particular behaviors. The experimental context represents those environmental characteristics that are designed by the psychologist to elicit a particular response or test score. The experiential and performance contexts are always nested within the ecological context, whereas the experimental context may or may not be contained within the first three contexts. Berry suggested that the degree to which the experimental context is nested within the other three contexts represents the ecological validity of the experimental task.

Four effects parallel the four contexts just described: achievements, behaviors, responses, and scores. Achievements refer to the complex, long-standing behavior patterns that develop as an adaptive response to the ecological context and include established and shared patterns of behavior that can either be found in an individual or distributed in a cultural group. Behaviors are those acts that have been learned over time in the experiential context and include the skills, traits, and attitudes that have been acquired in particular roles or fostered by specific training or education. Responses, on the other hand, are those 'fleeting' performances that occur in response to immediate stimulation or experience. Finally, scores are those behaviors that are observed, measured, and recorded during a psychological experiment or testing. Berry argued that between each of the four environmental contexts and effects exists a relationship that should be considered in any effective cross-cultural research.

The complexity of Berry's model may have limited its usefulness to some researchers. However, the general ecological approach has proven to be very useful in recent research investigating motivation, achievement, and related constructs. Specifically, such research has arguably produced data with acceptable ecological validity and has specified a range of experiential, performance, and experimental variables that are associated with motivation and achievement across a range of "natural" habitats.

2.3. Attribution Theory

The problem of developing an attributional model that might form a more adequate basis for cross-cultural research into achievement motivation has been a focus of older studies as well as more recent studies. These studies suggest that achievement motivation should be analyzed in terms of the subjective meaning of attributions (i.e., reasons given for success or failure) of a group or the persons who comprise that group. In other words, research into achievement motivation should be concerned with the phenomenology (i.e., definitions in context) of achievement attributions. This phenomenological approach to defining attributions cross-culturally requires that researchers elicit attributions of achievement behavior from people of different cultures rather than imposing inappropriate definitions on their achievement behavior.

In Western society, research indicates that assessments of achievement in terms of success and failure are best thought of as psychological reactions to outcomes and not as objective outcomes themselves. Thus, successes and failures are defined as much in terms of subjective feelings attached to given outcomes as in terms of the nature of the outcomes themselves. Moreover, attributions for success or failure that produce the greatest feelings of pleasure or regret are those related to ability (e.g., "I'm clever; therefore, I succeeded") and effort ("I tried hard; therefore, I succeeded"). More generally, attributions that enable a person to demonstrate to self or others that he or she possesses desirable qualities are considered to be important motivators. However, research concerned with examining the antecedents of success and failure across cultures suggests that ability and effort attributions are not necessarily salient to all groups. Hence, where research based on attribution theory emphasizes ability and effort as the most important perceived causes of perceived success and failure, it may be culture specific. For cross-cultural researchers, then, it is necessary to discover attributions that are salient to different cultural groups in eliciting feelings of success or failure.

2.4. Goal Theory

Goal theory developed by Ames, Dweck, Nicholls, Pintrich, and Midgley, provides a model that enables researchers to identify similar motives in diverse cultures, even if such behavior may vary in frequency and importance across cultures. Goal theory emphasizes the importance of goals and definitions of self in relationship to such goals. Goal theory differs from attribution theory in that attribution theory explores perceived reasons for success or failure while trying to succeed, whereas goal theory explores perceived reasons for trying to succeed

in the first place. These reasons for trying to succeed are labeled "goals" in the context of goal theory. Some important goals (or goal "orientations") postulated by goal theory are the following:

- *Ability-oriented (performance) motivation.* This approach is characterized by striving to maintain a favorable perception of one's ability to demonstrate competence. That is, the goal is to maximize the subjective probability of attributing high ability to oneself in culturally relevant areas. This approach, as with all goals in goal theory, focuses on the goal or function of behavior for the person rather than on the inevitably ambiguous external form of any given behavior itself.
- *Task-oriented (mastery) motivation.* This approach is characterized by behavior where the focus is on producing a product or solving a problem rather than on demonstrating one's competence. In contrast to ability-oriented behavior, during task-oriented behavior, people are not attributing outcomes to ability or seeking to demonstrate ability. The completion of the task itself is its own reward, and the development of mastery is perceived as intrinsically satisfying.
- *Socially oriented (social) motivation.* This approach is characterized by behavior where the goal is to maximize the probability of attributing high effort to oneself (and to minimize the probability of attributing low effort to oneself) in conformity with perceptions of the probability of gaining social approval, respect, or recognition. This orientation, which is really a class of orientations, is based on the assumption that effort, unlike ability, is voluntary and is something that anyone can display. Therefore, it can indicate conformity to social norms or virtuous intent, and this in turn can lead to social approval, respect, or recognition. Conversely, lack of effort in conforming to social norms or virtues may result in reduced approval, respect, or recognition and, hence, is to be avoided.

Goal theory has provided, and continues to provide, a cogent and coherent platform for investigating cross-cultural issues in achievement and achievement motivation, especially where the purposes for achievement are the focus of investigation. This is because goal theory allows for a range of possible intents (e.g., purposes, reasons) for achievement and for the possibility that the range and salience of possible intents may vary considerably from group to group. As a result, significant cultural differences in goal orientations, and their impact on motivation and achievement, have been noted in a range of studies. However, there is also contrary evidence from studies by McInerney and others suggesting that the goals that people espouse, and the effects of these goals on motivation and achievement, may be relatively stable across cultures.

3. APPLYING SOCIAL COGNITIVE CONCEPTUALIZATIONS OF MOTIVATION AND ACHIEVEMENT

In general, social cognitive theories of motivation, particularly goal theory and PI theory, have alerted researchers to the ways in which motivation and achievement may be differentially constructed across cultural (and other) contexts. However, a significant remaining question from this perspective is the following: Is there a single motivational environment that can be constructed in most applied settings (e.g., schools, hospitals, businesses, government organizations), especially when these settings are typically diverse in many ways with respect to individuals' ages, genders, cultural backgrounds, and socioeconomic status? This is an important question because a range of different motivational environments cannot be easily generated in most applied settings. Theory and research have gone a considerable distance in demonstrating that mastery goals are preferable to ability performance goals in a wide range of situations. This is because mastery goals focus on an individual's attention to the task(s) to be completed rather than on the individual's perceived ability (or lack thereof) to complete the task(s). Psychologically, this appears to "free" individuals to perform at optimal levels, even in the face of difficulties. Indeed, it might also be that performance goals are dysfunctional in many applied settings, although not in all cases. More recently, social goals have also been shown to have adaptive effects on an individual's engagement in a variety of achievement and social situations. These social goals may be particularly relevant in non-Western cultural settings, where the social dimensions of motivation and achievement may be more salient than in Western settings.

4. METHODOLOGICAL IMPROVEMENTS IN APPLIED SETTINGS

The preceding sections of this article have outlined (a) the need for a cross-cultural understanding of achievement and achievement motivation and (b) some relevant

conceptualizations and theories that may support research into achievement and achievement motivation from cross-cultural perspectives. This section deals with methods for researching these culturally variable constructs in appropriate ways in applied settings (e.g., schools, workplaces, sports settings).

It is obviously not acceptable to use "one size fits all" approaches when researching achievement and achievement motivation across cultures. Rather, culturally sensitive studies in achievement and achievement motivation are needed for the following purposes:

- To establish whether the behaviors and responses being examined across cultures are functionally, conceptually or metrically equivalent
- To establish that the constructs and tools used in research are culturally appropriate to all groups in the study
- To examine achievement and achievement motivation in "real life," as opposed to laboratory or similar situations, whenever possible
- To consider in detail, at both theoretical and methodological levels, the potential mediating processes that might intervene among culture, achievement motivation, and achievement
- Not to conceptualize achievement orientation solely as an individual personality-based characteristic
- Not to use cultural deficit and deprivation models (or, at least, not without justification) to explain the differential performance of various cultural groups
- Not to define any culture too narrowly and fail to recognize the range of experiences, perceptions, and behaviors represented in any given culture

Following these steps should help to avoid pseudoetic (e.g., Western individualist and competitive) ways in which to conceptualize achievement and achievement motivation in cultures where these values might not apply. At least these steps should contribute to a viable framework for examining achievement and achievement motivation cross-culturally in that they provide explicit procedures for developing culturally specific contents that may be applied in different cultural settings with comparable theoretical meanings.

4.1. Informal Assessment of Motivation in Applied Settings

Formal measurement of motivation (i.e., using psychometric instruments) occurs relatively rarely in applied settings such as schools and workplaces. However, informal assessments are made regularly in applied settings, particularly by those responsible for "performance management" (e.g., managers, supervisors, teachers). Informal indicators of motivation typically include a demonstrated willingness to be involved in assigned activities, the selection of appropriate activities where choice is available, the level of intensity or involvement in given activities, and the quality of performance outcomes.

These informal indicators of motivation are assessed or determined, to a large extent, by the frame of reference of the individual (typically a superior) making a determination. In other words, teachers, managers, and supervisors consider their subordinates to be more or less motivated in terms of the superiors' frames of reference. This process of evaluation, however, may cause problems in diverse settings (particularly in cross-cultural settings). For example, motivation in the workplace is often conceptualized by supervisors in terms of material or technical output (i.e., productivity). However, what happens when output is conceptualized (by subordinates) in "unconventional" ways, for example, by improving the relational climate of the workplace or classroom and by contributing to the work of others? In such cases, supervisors may consider subordinates to be "unmotivated" when the subordinates consider themselves to be highly motivated. A similar point could be made with respect to the balance between fulfilling social responsibilities (e.g., looking after family and friends) and fulfilling more conventional responsibilities (e.g., being on time for work, being available for after-work/after-school functions).

There are two major points being made here. First, motivation and the demonstration of motivated behavior may be conceptualized very differently by different cultural (or other) groups. Second, evaluations of motivation and motivated behavior are dependent on one's frame of reference. This adds considerable complexity to the evaluation of motivation in applied settings and argues against defining motivated behavior too narrowly, especially as the cultural diversity or other diversity of any setting increases.

5. CONCLUSION

This article has reviewed basic approaches to cross-cultural studies of achievement and achievement motivation, especially in applied settings such as schools, classrooms, and workplaces. It did so by highlighting some major limitations of the concepts and methods used to guide early studies while still recognizing the

value of this early work. It also presented a range of theoretical and methodological options that are currently being developed to enhance cross-cultural research in general and cross-cultural research into achievement and achievement motivation in particular.

A perennial problem for cross-cultural research is generating data that are universally comparable while also being culturally valid. Recent studies have demonstrated much more sophisticated and sensitive approaches to the issues of achievement and achievement motivation in cross-cultural settings.

See Also the Following Articles

Motives and Goals ■ Subjective Culture

Further Reading

Ames, C. (1984). Competitive, cooperative, and individualistic goal structures: A motivational analysis. In R. Ames, & C. Ames (Eds.), *Research on motivation in education*, Vol. 1: *Student motivation* (pp. 177–207). San Diego: Academic Press.

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, *84*, 261–271.

Atkinson, J. W., & Feather, N. T. (Eds.). (1966). *A theory of achievement motivation*. London: Wiley.

Atkinson, J. W., & Raynor, J. O. (1974). *Motivation and achievement*. Washington, DC: V.H. Winston.

Berry, J. W. (1980). Ecological analyses for cross-cultural psychology. In N. Warren (Ed.), *Studies in cross-cultural psychology* (Vol. 2, pp. 157–189). London: Academic Press.

Bright, J. A. (1994). Beliefs in action: Family contributions to African-American student success. *Equity and Choice*, *10*, 5–13.

Conchas, G. Q. (2001). Structuring failure and success: Understanding the variability in Latino school engagement. *Harvard Educational Review*, *71*, 475–504.

Conner, R. F. (1998). Toward a social ecological view of evaluation use. *American Journal of Evaluation*, *19*, 237–241.

Gomez, M. J., & Fassinger, R. E. (1994). An initial model of Latina achievement: Acculturation, biculturalism, and achieving styles. *Journal of Counseling Psychology*, *41*, 205–215.

Graham, S., & Taylor, A. (2002). Ethnicity, gender, and the development of achievement values. In A. Wigfield, & J. Eccles (Eds.), *The development of achievement motivation* (pp. 121–146). San Diego: Academic Press.

Graham, S., Weiner, B., & Benesh-Weiner, M. (1995). An attributional analysis of the development of excuse-giving in aggressive and nonaggressive African American boys. *Developmental Psychology*, *31*, 274–284.

Johnson, G. M. (1994). An ecological framework for conceptualizing educational risk. *Urban Education*, *29*, 34–49.

Kaplan, A., & Maehr, M. L. (1999). Enhancing the motivation of African American students: An achievement goal theory perspective. *Journal of Negro Education*, *68*, 23–41.

Kornadt, H. J., Eckensberger, L. H., and Emminghaus, W. B. (1980). Cross-cultural research on motivation and its contribution to a general theory of motivation. In H. C. Triandis & W. Lonner (Eds.), *Handbook of cross-cultural psychology* (Vol. 3, pp. 223–320). Boston: Allyn & Bacon.

Maehr, M. L., & Braskamp, L. A. (1986). *The motivation factor: A theory of personal investment*. Lexington, MA: Lexington Books.

Maehr, M. L. (1974). Culture and achievement motivation. *American Psychologist*, *29*, 887–896.

McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. (1953). *The achievement motive*. New York: Appleton–Century–Crofts.

McClelland, D. C., Baldwin, A. L., Bronfenbrenner, U., & Strodbeck, F. L. (1958). *Talent and society: New perspectives in the identification of talent*. Princeton, NJ: Van Nostrand.

McInerney, D. M. (1994). Psychometric perspectives on school motivation and culture. In E. Thomas (Ed.), *International perspectives on culture and schooling*. London: London University, Institute of Education.

McInerney, D. M. (1995). Goal theory and indigenous minority school motivation: Relevance and application. In P. R. Pintrich, & M. L. Maehr (Eds.), *Advances in motivation and achievement* (Vol. 3, pp. 153–181). Greenwich, CT: JAI.

McInerney, D. M., Roche, L., McInerney, V., & Marsh, H. W. (1997). Cultural perspectives on school motivation: The relevance and application of goal theory. *American Educational Research Journal*, *34*, 207–236.

McInerney, D. M., & Swisher, K. G. (1995). Exploring Navajo motivation in school settings. *Journal of American Indian Education*, *34*(3), 28–51.

McInerney, D. M., Yeung, S. Y., & McInerney, V. (2001). Cross-cultural validation of the Inventory of School Motivation (ISM). *Journal of Applied Measurement*, *2*, 134–152.

Moeller, A., & Kramer, E. (1995). The teaching of culture and language in the second language classroom: Focus on the learner. *International Journal of Educational Research*, *23*, 571–652.

Neuman, S. B., & Celano, D. (2001). Access to print in low-income and middle-income communities: An ecological study of four neighborhoods. *Reading Research Quarterly*, *36*(1), 8–26.

Niles, S. (1998). Achievement goals and means: A cultural comparison. *Journal of Cross-Cultural Psychology*, *29*, 656–667.

Ogbu, J. U., & Simons, H. D. (1998). Voluntary and involuntary minorities: A cultural–ecological theory of school

performance with some implications for education. *Anthropology & Education Quarterly*, *29*, 55–88.

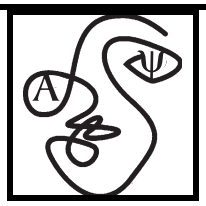
Salili, F., & Hoosain, R. (2002). Cross-cultural differences in affective meaning of achievement: A semantic differential study. In D. M. McInerney, & S. Van Etten (Eds.), *Research on sociocultural influences on motivation and learning* (Vol.2). Greenwich, CT: Information Age.

Strohschneider, S., & Guss, D. (1998). Planning and problem solving: Differences between Brazilian and German students. *Journal of Cross-Cultural Psychology*, *29*, 695–716.

Villani, C. J. (1999). Community culture and school climate. *School Community Journal*, *9*, 103–105.

Watkins, D., McInerney, D. M., Lee, C., Akande, A., & Regmi, M. (2002). Motivation and learning strategies. A cross-cultural perspective. In D. M. McInerney, & S. Van Etten (Eds.), *Research on sociocultural influences on motivation and learning* (Vol. 2). Greenwich, CT: Information Age.

Weiner, B. (1994). Integrating social and personal theories of achievement striving. *Review of Educational Research*, *64*, 557–573.



Motivation to Diet and Exercise

Larry M. Leith
University of Toronto, Toronto, Ontario, Canada

1. Introduction
2. Motivational Considerations in Diet and Exercise
3. Individual Intervention Strategies to Improve Diet and Exercise
4. Concluding Remarks
Further Reading

GLOSSARY

behavior therapy A field of psychology that provides information telling people how to change their individual habits; it involves self-control to make people's lives better.

motivation Term used to explain why people do, or do not do, certain behaviors.

obesity When an individual has a body mass index (BMI) of at least 30.

protection motivation theory Theory suggesting that people are motivated to protect themselves from harmful or stressful life events.

self-determination theory Viewpoint evaluating the relative importance of intrinsic and extrinsic rewards for human behavior.

self-efficacy Term referring to situation-specific self-confidence.

People are continually trying to exercise more and eat diets that are conducive to good health. The problem of childhood and adolescent obesity continues to be a major focus of concern for health care professionals in North America. This article describes the necessary tools to make lifestyle changes with the potential to reverse this troublesome trend.

1. INTRODUCTION

The problems associated with obesity and chronic degenerative diseases, such as heart disease, arteriosclerosis, hypertension, osteoarthritis, and type 2 diabetes, are now reaching epidemic proportions. The prevalence of obesity, where an individual has a body mass index (BMI) of at least 30, is increasing worldwide. Currently, there are approximately 300 million obese individuals around the world. More alarmingly, in the United States, approximately 30% of all adults are categorized as obese. Despite this prevalence being somewhat lower in Canada, the rate has increased significantly there during recent years as well. In 2003, Bouchard reported that obesity costs the unregulated U.S. health care system approximately $100 billion annually. Therefore, it is time to take a serious look at how to reverse this troublesome trend.

In 2000, data collected by the U.S. Department of Health revealed that nearly 30% of the population is physically inactive. It is interesting to note how this percentage corresponds exactly with the percentage of individuals who are considered to be obese. Also, an additional 40 to 50% of the U.S. population does not engage in the necessary amount of physical activity to achieve health benefits. Through a process known as tracking, these statistics are reported to have remained

©2004 Elsevier Inc.
All rights reserved.

relatively consistent for the past 10 years or so. This is unfortunate given that frequent physical activity has been shown to benefit every system of the human body. For example, exercise physiologists have documented several important changes in the body that accompany participation in an ongoing exercise program. These include a decreased chance of heart disease and stroke, an increase in good (HDL) cholesterol, a decrease in bad (LDL) cholesterol, a lowered risk of osteoporosis (i.e., brittle bones), a decreased percentage of body fat, an increased lean muscle mass, lower blood pressure, and an improved immune system. In addition, in 1998, Leith reported that regular participation in physical activity has been found to be associated with several psychological benefits such as lowered anxiety, lowered depression, improved self-concept, and better overall mood.

In 1996, Manley, the former U.S. surgeon general, made the following observation. Because physical activity is so effective in preventing diseases, preventing premature death, and promoting a high quality of life, it should be accorded the same level of attention that is given to other important health practices that affect the nation. This assertion suggests that there is absolutely no doubt that physical activity can be described as efficacious in the prevention and treatment of many medical conditions, including obesity. However, despite this and extant research highlighting the benefits of physical activity and a healthy diet, people remain inactive and eat poorly. Therefore, it is important to consider the "why" of this phenomenon. The next section summarizes contemporary motivational theories relating to these topics and examines their relevance for influencing the health behaviors of diet and exercise.

2. MOTIVATIONAL CONSIDERATIONS IN DIET AND EXERCISE

The term "motivation" is actually a derivative of the Latin word *movere,* meaning "to move." It is the psychological term used to explain why people do, or do not do, certain things. For example, why do some individuals exercise regularly and eat healthy diets, whereas others choose the options of inactivity and eating junk foods? This is not an easy question to answer, but the bottom line is that people's underlying motivation can be understood by looking at three characteristics of behavior. First, people's selection of certain activities over others indicates something about their motivation. To paraphrase, what is it that is important to people personally? Their answers to that question will indicate their likelihood or reluctance to engage in a specific behavior. Second, the effort that people put into any activity also reveals their underlying motivation. In other words, the more effort people put into any behavior, the more motivated they are to accomplish that goal. The harder people work toward accomplishing a task, the more they want it. Finally, the degree of persistence that people demonstrate in their pursuit of a goal also reveals their level of motivation. If they refuse to give up and really want it but still give up easily, they did not want it enough in the first place. Therefore, it is important that each and every individual is aware of these three benchmarks of motivational theory.

Now that we have examined the basic components of motivation, we look at theories of motivation that have specific relevance for dieting and exercise. It is far beyond the scope of this article to cover the whole range of motivational theories. In 2003, Carron and colleagues published a book summarizing a wide range of motivational theories. This section focuses only on self-efficacy theory, protection motivation theory, and self-determination theory. In each case, a summary of the theory is provided with specific implications for diet and exercise.

2.1. Self-Efficacy Theory

The term "self-efficacy" can be summarized with an individual asking himself or herself the following two questions: "Am I capable of doing this activity?" and "Will I be successful in achieving my goals if I engage in this behavior?" The internal question specific to this article would be "Am I capable of successfully participating in an ongoing diet and exercise program?" Self-efficacy can also be described as situation-specific self-confidence.

In 1997, Bandura reported that an individual's concept of self-efficacy is molded by his or her past experiences, vicarious experiences, verbal persuasion, and physiological/emotional states. To summarize this theory in light of the article's objectives, self-efficacy can be enhanced in a variety of ways. First, people should remember to take "baby steps." In other words, they should start an exercise program slowly and try to improve their overall diet in small stages. Most people fail because they try to do too much too soon. Second, people should take advantage of watching demonstrations, television shows, and videos that show other people succeeding in their efforts. These options will improve people's belief in their ability to succeed in

their quest. Third, people should learn everything they can about the exercise and diet topics. More information is always better. They should not be shy about soliciting support from their significant others; doing so can go a long way toward improving their chances of success. Finally, people should be aware of how their physiological/emotional states can influence their resolve to diet or exercise. It is only natural to question the increased heart rate associated with exercise and/or the effect of mood on exercise or diet compliance. People should become familiar with these outcomes and recognize them as harbingers of success in their future efforts.

2.2. Protection Motivation Theory

This theory suggests that people are motivated to protect themselves from harmful or stressful life events. It also considers coping strategies for these occurrences. To summarize, the protection motivation theory maintains that an individual's motivation to engage in a health-related behavior is influenced by two considerations. First, threat appraisal represents a conscious evaluation of a person's perceived vulnerability and the possible severity of a medical problem. The internal questions ask "Am I at risk, and how serious are the possible outcomes of not exercising and of eating a poor diet?" Second, coping appraisal reflects an individual's perception of whether he or she will be successful in implementing the appropriate behavior and, if so, whether these actions will remove the threat. The internal questions ask "Am I capable of engaging in regular exercise and maintaining a healthy diet and, if so, will I be less at risk for developing chronic degenerative diseases?" The answers to these questions will provide a good indication of the person's intention to improve exercise and diet habits. Intention remains a critical ingredient in predicting human behavior.

In 2000, Floyd and colleagues and Milne and colleagues published meta-analytic reviews suggesting the efficacy of the protection motivation theory in predicting health behaviors. The former study examined 65 published articles with 29,650 participants and concluded that increased threat severity, threat vulnerability, self-efficacy, and response efficacy resulted in facilitated health behaviors. The latter study reviewed 29 studies with 7700 participants and also concluded that the protection motivation theory is predictive of facilitated adaptive intentions and improved health behaviors. Therefore, it appears that an individual's perceived severity of threat and vulnerability from inaction, as well as the chances of being successful in changing these health behaviors, will seriously affect the desire to exercise more and to maintain a healthy weight.

2.3. Self-Determination Theory

This theory evolved in an attempt to understand the relative importance of intrinsic and extrinsic rewards for human behavior. Originally, it was believed that if intrinsic rewards were present, extrinsic rewards would be absent, but empirical research disproved this hypothesis.

Deci's 1992 development of self-determination theory attempted to address the discrepancy. This approach maintains that extrinsic and intrinsic motivations exist on a continuum. At one end of the continuum is amotivation, indicating the complete absence of motivation toward engaging in a particular behavior such as exercising or losing weight. At this point on the continuum, the person has no desire to change his or her lifestyle. Extrinsic motivation lies in the middle of the continuum and is multidimensional in nature. It consists of three additional dimensions. Consider the scenario where a family doctor tells a patient that he or she must lose weight and start exercising to avoid serious medical problems. In this case, the patient may try to follow the doctor's advice simply to avoid illness. Slightly farther along the continuum, the patient may lose enough weight and exercise to no longer be considered a health risk, and the patient continues the program only because he or she believes that it must be done. Finally, the individual may continue to "count calories" and exercise three times per week, but only because doing so is necessary to avoid illness.

At the farthest point on the continuum lies intrinsic motivation, and it is also multidimensional. This type of motivation occurs when a person performs a behavior for the sole pleasure it provides. Intrinsic motivation is also a multidimensional construct. For example, acquiring the knowledge of how to count calories effectively, eat a more healthy diet, and perform a new exercise represents the first stage of intrinsic motivation. During the second stage, the satisfaction of losing weight and becoming a regular exerciser becomes more important. Finally, once an individual reaches the intrinsic motivation point on the continuum, he or she will continue to exercise solely for the pleasant sensations associated with the physical activity. Likewise, the person will continue to maintain a healthy weight simply because it feels good to do so.

Now that we have examined the theories of motivation most relevant to diet and exercise, it is time to learn how to apply this important knowledge. The next

section discusses the behavioral tools that people need to take control of their lifestyle choices.

3. INDIVIDUAL INTERVENTION STRATEGIES TO IMPROVE DIET AND EXERCISE

The field of behavior therapy provides important information concerning the making or breaking of individual habits. Practitioners who endorse this approach maintain that people often fail in their self-control efforts because they do not consider environmental factors or the "ABCs of behavior." Environmental stimuli (i.e., cues) that affect people's behavior represent the As. The behaviors themselves are termed the Bs, and the rewards or punishments that closely follow the behaviors are labeled the Cs. The following specific interventions use this ABCs approach to behavior change.

3.1. Strategy 1: Removing Potent Cues That Are Associated with a Problem Behavior

One serious handicap to the average weight control program is the typical pairing of eating and snacks that are readily visible. Not many people can walk by a bowl of chips or a box of chocolates without sampling these treasures. For this reason, if people are attempting to lose weight, they should not have the fattening snacks showcased in a well-traveled area of the house. In terms of exercise, many individuals report that televisions and computers are major impediments to being physically active. If this is a problem for people, they should make a deal with themselves not to walk by their televisions or computers until they have completed their exercise sessions. This will eliminate the automatic pairing of watching television or surfing the Internet with inactivity.

3.2. Strategy 2: Making the Problem Behavior Unlikely or Impossible

Following from the previous example, the simplest way not to indulge in high-calorie snacks is simply not to have them in the house. This way, if people feel that they are weakening, it takes a good deal of effort to make a special trip to the store, so they are less likely to do so. When people shop for groceries, they should not buy these fattening foods. Similarly, if people find that they fail to exercise on days when they go to a tavern or to a friend's place, they should make an agreement with themselves to avoid these venues on exercise days. People will still have many opportunities to enjoy these situations that are getting in the way of their exercise programs.

3.3. Strategy 3: Developing New Behaviors to Replace Maladaptive Ones

This intervention invites people to identify the existing connection between the usual cues to action and the problem behaviors. For example, if people find that they tend to overindulge in snacking when they are feeling angry or lonely, they should consider an alternative behavior. Many individuals report that writing down their thoughts or feelings on a pad of paper can be therapeutic. If people are trying to lose weight, this is a better option than snacking. Taking this example to the next level, people might want to consider going for a walk or going to a fitness facility to deal with their feelings. Exercise has been shown to improve mood, and exercising in a social setting can help to alleviate feelings of anger or loneliness.

3.4. Strategy 4: Developing Cues That Will Foster a Desired Behavior

People should think in terms of using environmental cues that can work for them in pursuit of their goals. Individuals who want to lose weight should consider taping pictures of their dream tropical island getaway destinations on the fronts of their refrigerators. This way, each time they are tempted to snack, they can visualize themselves looking good in their swimwear on the beach. This will give them the needed strength to resist temptation. People who want to exercise regularly should consider placing their walking/running shoes and/or sweat suit at the point of entry into their residences or on their bedposts. This way, the first thing that they see when they get home from work or wake up in the morning is exercise equipment. This will prove to be a strong reminder for them to walk and/or jog.

3.5. Strategy 5: Rewarding Oneself for Exhibiting a Good Behavior

The whole idea of whether or not people should reward themselves is contingent on the principle of their success in achieving the target behavior. In other words, if

people maintain their daily caloric intake or achieve their exercise goals, they need to give themselves a "pat on the back." It is important that this reward closely follows the desired behavior. Effective self-control requires a reorientation of the way in which people reward themselves. If they limit themselves to a certain number of calories or engage in their exercise programs on any given day, they are entitled to a reward. People should take this reward sooner rather than later because they deserve it and it will lead to greater compliance.

3.6. Strategy 6: Rewarding Oneself with What One Really Wants to Do Anyway

If there is a behavior that people really like to do on any given day, such as watching television or surfing the Internet, they should use one or both of these activities as potential rewards for achieving their daily goal. So, if people like to watch a favorite television show or use their computers, they should allow themselves to do so only after they have gone for their walks or reduced their caloric intake for that day. This represents a way in which people can reward themselves with things they have already identified as important to them personally.

3.7. Strategy 7: Using a Point–Reward System to Motivate Oneself

Sometimes, achieving a goal can seem a long way off. For this reason, it is important for people to use a point–reward system so as to continue to motivate themselves. For example, if people's goal is to lose 20 pounds or to be able to jog 2 miles, it is important to remember that these goals can be reached only with "baby steps." This means that people should set smaller, short-term goals to achieve their desired results. Losing 1 pound per week or walking/jogging a certain distance during that week allows people to give themselves a "point" or "star" for their accomplishments. People should record these points or stars in diaries or on their monthly calendars as a form of self-monitoring. They should make a deal with themselves that once they have accumulated a certain number of points or stars, they can trade these in for something they really want such as a night on the town, a new video, or a CD that was recently advertised. They should continue to set new reward limits and to treat themselves every time they have achieved their short-term goals.

3.8. Strategy 8: Being Creative in Finding New Ways in Which to Reward Oneself

One of the most often overlooked sources of reinforcement is of a social nature, for example, people's friends, families, and coworkers. People should not be shy about enlisting their families, friends, and coworkers for support. A pat on the back or a "thumbs up" from a significant other can go a long way toward helping people to achieve their behavioral goals of losing weight and exercising more.

3.9. Strategy 9: Using a Behavioral Contract to Help Achieve One's Goals

Writing up an exercise and/or diet contract can be a very effective strategy. It requires people to put in writing their diet or exercise aspirations. For example, the number of days per week people are going to exercise and the number of daily calories they are going to consume are formalized with the written contract. It is also a good idea to include the mention of a reward that people will receive if successful as well as a response cost if they slack off in their mission. People sign the contract and have significant others also sign the document. This will increase the chances that they will be successful in their efforts.

4. CONCLUDING REMARKS

This article has examined the problems of increasing obesity trends and declining exercise habits in the North American population. Although these problems are indeed factual, it is important to put both topics into a healthy perspective. This article is not intended to suggest that everyone should become obsessive–compulsive about their diets and exercise patterns. The best approach is for people to find a happy medium where they exercise moderately and eat a healthy diet that is low in fats and includes lots of fruits and vegetables. If people buy into this philosophy, the intervention strategies outlined in the preceding section will provide them with the behavioral tools needed to achieve healthier lifestyles. Most experts would suggest that behavior can be changed within a matter of weeks using the behavior therapy approach.

See Also the Following Articles

Eating, Psychology of ■ Health and Culture ■ Health Psychology, Cross-Cultural ■ Obesity

Further Reading

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Bouchard, C. (2003, September). *Alarming facts and potential causes of the growing obesity epidemic*. Paper presented at the "Obesity: Weighing the Options" symposium, Toronto, Canada.

Carron, A. V., Hausenblas, H. A., & Estabrooks, P. A. (2003). *The psychology of physical activity*. New York: McGraw–Hill.

Deci, E. I. (1992). On the nature and function of motivation theories. *Psychological Science*, *3*, 167–171.

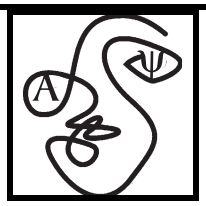
Floyd, D. L., Prentice-Dunn, S., & Rogers, R. W. (2000). A meta-analysis of research on protection motivation theory. *Journal of Applied Social Psychology*, *30*, 407–429.

Leith, L. M. (1998). *Exercising your way to better mental health*. Morgantown, WV: Fitness Information Technology.

Manley, A. E. (1996). Preface. In U.S. Department of Health and Human Services, *Physical activity and health: A report of the surgeon general executive summary*. Washington, DC: U.S. Department of Health and Human Services.

Milne, S., Sheeran, P., & Orbell, S. (2000). Prediction and intervention in health related behavior: A meta-analytic review of protection motivation theory. *Journal of Applied Social Psychology*, *30*, 106–143.

U.S. Department of Health and Human Services. (2000). Physical activity and fitness. In *Healthy People 2010*. www.health.gov.



Motives and Goals

Uwe E. Kleinbeck
University of Dortmund, Bochum, Germany

1. Introduction
2. Goal Effects: Mechanisms and Moderators
3. Developing a Personal Goal
4. Acceptance of Goals
5. How to Pursue Goals That Are (Still) Unknown
6. Managing Successful Goal-Oriented Work Behavior: The Participatory Productivity Management
7. Competencies in Managing Goal-Directed Behavior
8. An Integrative Model for the Effects of Motives and Goals
Further Reading

GLOSSARY

competencies Capabilities needed by leaders for supporting goal setting, translating goals into action, committing to goals, and dealing with failure in managing human resources at work in a goal-directed way.

dealing with failure Dealing with not reaching a goal; failure can lead to strengthened approach tendencies with a higher motivation to reach the goal, but it can also result in withdrawal from the task.

goal acceptance When acting persons accept goals set by others; if such persons experience positive influence potentials, they have the chance of negotiating new goals.

goal commitment The commitment that keeps people persistent in striving for goals; the degree of goal commitment is dependent on changes in current motivation.

management system Instruments that are able to measure performance, offer feedback, and stimulate participative goal setting to help people reach organizational goals.

moderators Variables that moderate the translation of goals into actions; the most important ones are task complexity, goal commitment, and individual ability.

motivating potentials Opportunities in a situation to reach the goals of a motive; the interaction of motivating potentials with motives determines work motivation.

Motives and goals work together to control the intensity, direction, and persistence of behavior at work. Motives are defined as classes of goals that determine, in interaction with motivating potentials of the work situation, the kind of goals people use to control their actions. Goal characteristics, mechanisms, and moderators of the translating process are the determinants of behavioral outcomes. Motives also play a role during this action phase. They affect the degree of goal commitment that people show while performing. Especially interesting is the impact of motives on the way in which people deal with failure. People differ in their ways of reacting to failure. Some are better equipped to deal with failure because they tend to increase motivation, stay with their goal, and finally reach it. Others withdraw from the task.

1. INTRODUCTION

Goals are the basic concepts used to understand and describe the direction and persistence of behavior at

©2004 Elsevier Inc.
All rights reserved.

work. Goals are effective on various levels of abstractness, with very specific goals controlling specific behaviors (e.g., motor acts). For example, a person can have the goal to press a button 300 milliseconds after a signal appears on a screen. On a more abstract level, goals are aggregated and often are called motives. A motive is defined as a highly aggregated class of goals, with variants of a basic theme being common to all of them. For example, a set of social goals forms a social motive acting as a stable disposition in persons across situations and over long time periods. Examples of other motives when the focus is on the development of personal goals are provided later in this article. In this context, motives are introduced as central determinants of personal goal setting in concrete situational contexts.

Goals can be set by an acting person internally or can be set externally by others. In the following, the focus is on internally set goals. The acceptance of externally set goals is described later in the article. Next, several functional properties of goals and three phases of goal-oriented action processes are discussed, and how goals and motives interact during these phases of action is discussed throughout the article.

Setting goals for oneself means anticipating expected outcomes of action that the person wants to reach. Goals also function as motivational states, directing behavior, influencing commitment to those goals, and keeping action more or less persistent toward those goals. Moreover, goals serve as reference values in a feedback loop in that they are used as standards to compare behavioral states in the course of actions with the intended goals. Finally, goals trigger the forming of emotions such as pride (in the case of reaching goals) and shame (in the case of failing to reach goals).

2. GOAL EFFECTS: MECHANISMS AND MODERATORS

According to Locke and Latham, goals determine the effort that people invest in their actions, goals direct behavioral acts toward the intended outcomes, and goals are responsible for the degree of persistence in keeping actions goal oriented. These are the mechanisms by which goals affect performance and outcomes. The effect of goals is moderated by a set of variables, and in this context the three most important variables are goal commitment, task complexity, and ability (Fig. 1).

Before going into details of the moderator effects, some characteristics of goals influencing their effect on performance should be noted. Hundreds of studies have shown that setting a specific difficult goal leads to higher performance than does a general intention to do one's best. This is true for individuals who have the ability necessary to perform a task. For people who do not have the necessary knowledge to perform a task effectively, setting a high outcome goal can have detrimental effects on performance.

In 2001, Seijts and Latham showed that in the case where people do not have the necessary knowledge to

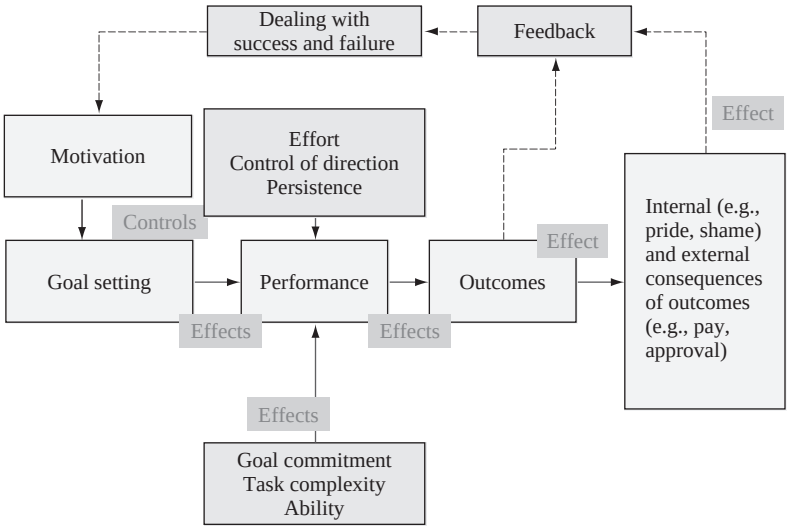


FIGURE 1 Concepts for understanding the translation of goals into action.

perform a task effectively, it is helpful to set a learning goal instead of an outcome goal. They defined a high learning goal as specific and difficult in terms of the number of strategies to be discovered to learn how to perform a task. In their study, they showed that a specific difficult learning goal leads to higher performance than does urging people to do their best on a task where they initially lack the requisite knowledge to perform it. Specific difficult outcome goals should be avoided because they appear to have a detrimental effect on performance during the learning process. The practical implication of these findings is that practitioners should focus on ways in which to increase workers' self-efficacy on tasks that the workers perceive as difficult and complex. This helps to avoid failure experiences with detrimental consequences for further goal setting and performance, and it also reduces negative effects on self-esteem.

Two additional variables in this model deserve some attention: goal commitment and the way in which actors deal with failure. Goal commitment is important to keep actions going so as to reach a goal. Highly committed persons stay with goals persistently, even when they encounter difficulties and hindrances along the way. They keep focused on the current task and resist temptations raised by other motivating potentials or motives.

Commitment to a goal actually controlling ongoing behavior directly depends on the motivation while pursuing the goal (Fig. 2). If there is no competitive motivational state, a persistent striving toward the original goal is not threatened at all. Although the strength of goal commitment for the ongoing action is affected by personal motives and situational characteristics when starting the action, there may be changes in both groups of variables that are responsible for either a decrease in goal commitment or an increase in motivation for a different action. This could lead to an early change in activity (due to changing goals) before the originally intended goal has been reached. Other determinants of a high goal commitment are high self-efficacy, a high expectancy to reach the goal, and a high instrumentality of the outcome aimed for higher order outcomes in the future. A success-oriented achievement motive could also be helpful for keeping goal commitment high if the motivating potential of work is high with respect to the achievement motive system.

Setting high goals at work increases the probability of high performance. But high goals also carry the risk of failure. Although occasional failure should not be dangerous for successful work, it can be dangerous in the long run, especially when it has a negative influence on the self-esteem of acting persons. In the positive case, failure strengthens approach tendencies and goal commitment. The original goal stays as high as it was before the failure, or it is set even higher to invest more effort than before. This increases the chances of reaching the goal (Fig. 3). In the negative case, failure urges people to withdraw from the task, goal commitment goes down, and the acting persons give up on original goal. This leads to lower performance.

A large amount of research clearly shows the convincing effects of goals on behavior at work. Not only can performance be increased by high and specific

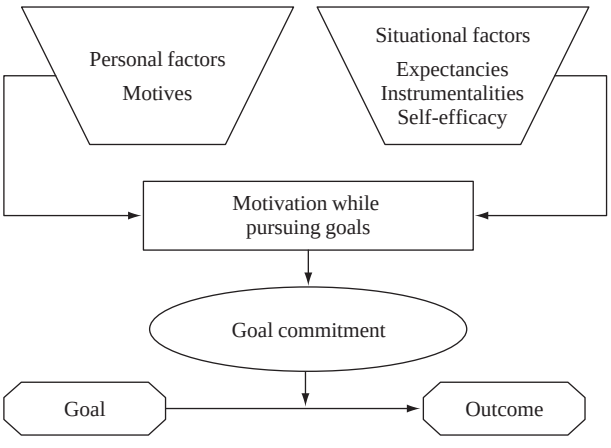


FIGURE 2 Determinants of goal commitment.

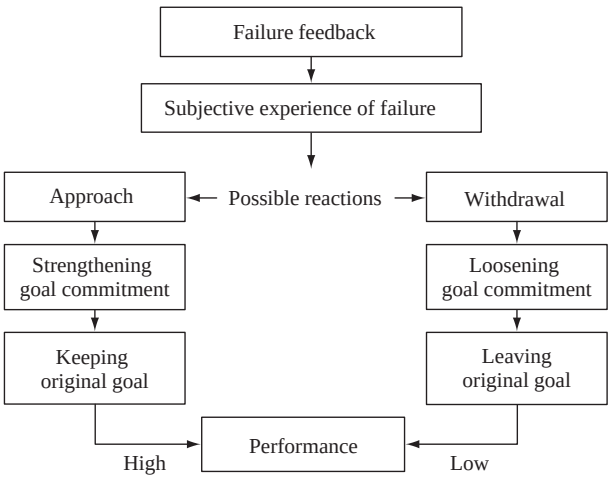


FIGURE 3 Possible reactions to failure: Approach or withdrawal.

outcome or learning goals, but withdrawal behaviors, such as absences and fluctuations, also can be influenced by goals. Important moderators are goal complexity, goal commitment, and the competence to deal with failure in an adequate mode. Motives are especially effective as central determinants of goal commitment. Of course, they unfold their influence on goal commitment in interactions with motivating potentials, changing with various situational aspects occurring in the course of an action.

3. DEVELOPING A PERSONAL GOAL

Whether a person sets high or low goals, whether the person directs his or her behavior in this or that direction, and whether the person persists more or less with one goal depend on the interaction of personal and situational factors. In motivational theory, these factors are called motives as personal factors and motivating potentials on the side of the situational factors. As stated previously, motives are highly aggregated classes of goals, and motivating potentials provide the occasions for realizing motive goals. Important motives in modern industrial and service-oriented societies include the following:

- Curiosity
- Achievement
- Affiliation
- Power
- Aggression
- Hunger, thirst, and sexual motives

For each motive, there are motivating potentials in the situation in which persons are acting. Motivating potentials for persons with a high curiosity motive are high when a situation enables them to increase their knowledge. For more information about motivating potentials related to various motives, see Table I.

Motives and motivating potentials interact in determining goal direction and goal difficulty (Fig. 4). Empirical evidence has been reported for several motive domains. For purposes of illustration, the interaction of the achievement motive with the corresponding motivating potential of the work situation is described. The achievement motive is defined as striving for competing with a standard of excellence. Persons with a high achievement motivation like to know how efficient they are and how to improve their efficacy. To set high goals, these persons need challenging tasks that give them the opportunity to see for themselves how efficient they are and that allow them to find ways in which to increase their efficacy. Such tasks are characterized by a variety of task components. They are important, offer a certain amount of autonomy, and allow continuous feedback during task performance together with clear outcome feedback. In 1976, Hackman and Oldham referred to the combination of these factors as "motivating potential of tasks."

Outcome and performance feedback can be used by the acting person to calculate expectancies about the chances of reaching a goal. In the case of achievement motivation, expectancies are also important for the strength of motivating potentials because it makes sense for those with a high achievement motive to work with tasks that give them at least a moderate probability of success. In this way, it is possible for persons with a high achievement motive to learn as much as possible about their own efficacy.

TABLE I
Important Motives and the Corresponding Motivating Potentials of Situations

Motive	*Motivating potentials offer an opportunity to:*
Curiosity	Get better knowledge of what is going on
Achievement	Enhance one's competence and feel pride about having reached a goal
Affiliation	Build a new social relation and keep it in a good state
Power	Influence the behavior of others
Aggression	Perform harmful acts toward others
Hunger, thirst, and sexual motives	Get food and water and get a sexual partner

4. ACCEPTANCE OF GOALS

Very often in their lives, especially at work, persons are asked to accept goals that have not been set by themselves. In such cases, the question arises as to whether they accept these external goals for controlling and regulating their own behavior. Reacting on externally set goals, a person has at least three choices: accepting the goals, internalizing the goals, or using the goals to control goal-oriented actions (Fig. 5). This happens, of course, when the externally set goals are not in conflict

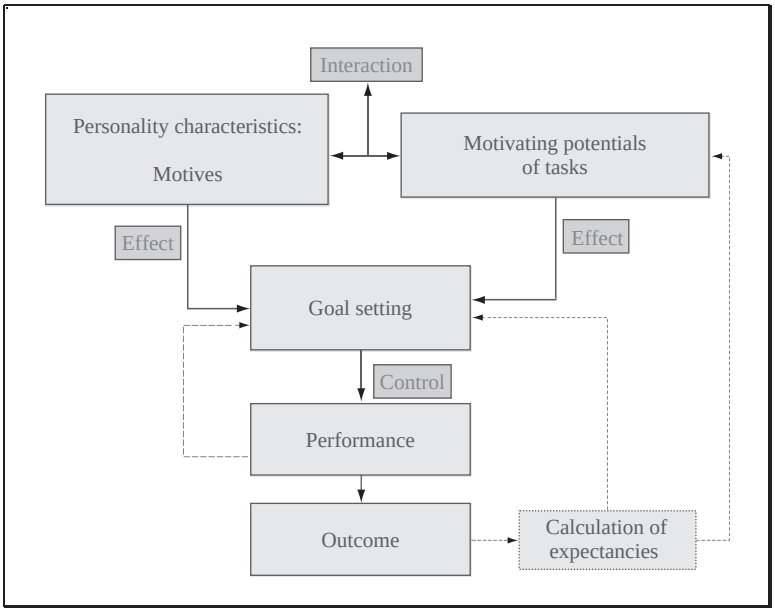


FIGURE 4 The formation of goals as the result of an interaction between personal and situational characteristics.

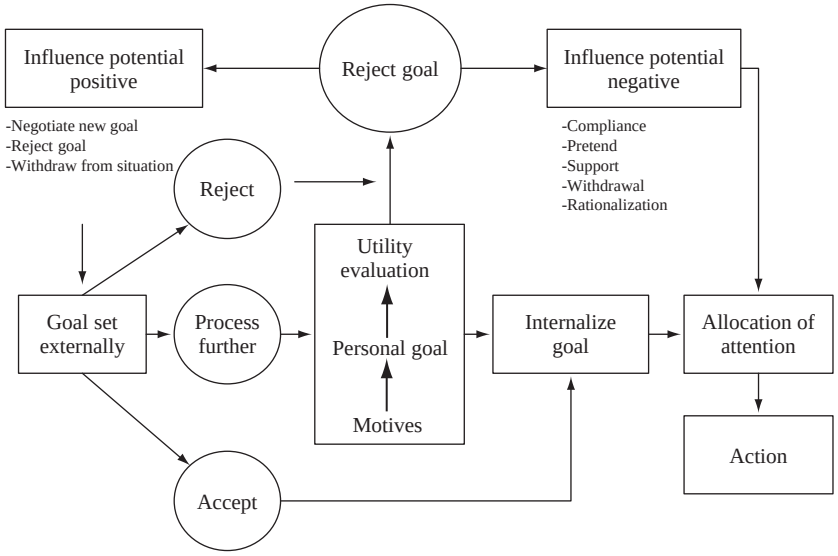


FIGURE 5 Concepts used to describe the process of accepting and internalizing or rejecting externally set goals. Adapted from Earley and Shalley (1991).

with personal goals, when they might even lead to outcomes with high personal valence, or when the costs of a conflict between personal and organizational goals is compensated by a stimulating reward.

On the other hand, one can reject goals directly when it appears clear from the beginning that there will be conflicts with personal goals and no compensation is in sight. In this case, the person begins to think about the possibility of influencing the leader who set the goal. If influence appears to be possible, negotiation is started to change the original goal, possibly leading to a new goal. If the influence potential of the situation is negative, the person must decide either to comply and accept the goal or to withdraw from the situation.

These considerations have some practical implications for motivating people at work. When setting goals externally, it makes sense to be aware of the motives of workers and employees because these motives are of essential importance to their personal goals. If managers know these goals, they can calculate in advance what effects their externally set goals will have. If conflict between externally set goals and personal goals is expected, managers either can reformulate their goals to minimize conflict or can use other techniques that make clear how important it is to work on the goals set by them even if such goals are not in line with personal goals. One technique could be to introduce a participative goal-setting program, as discussed later in the article.

5. HOW TO PURSUE GOALS THAT ARE (STILL) UNKNOWN

Activities at work in general are goal oriented. Often, goals are specific and difficult, leading to high performance. Goal-setting theory offers evidence for understanding these effects. But nowadays, more and more work activities are no longer related to clear and specific goals because the specific goals are not yet known when the activities are started. One classic example is given with design tasks. They are characterized by the fact that outcomes are not at all clear for a long time during work on the tasks. They are similar to the learning tasks mentioned earlier in the article. In design tasks, people cannot do the following:

- Analyze the conditions of action necessary for reaching the goals exactly.
- Run trial actions.
- Form language terms to describe actions.
- Coordinate component goals toward ultimate goals.
- Compare outcomes with goals stored in memory.

Although these activities cannot be regulated by specific goals as anticipated outcomes, design actions are not without direction given that the outcomes must be developed first. During activities of this kind, people must do the following:

- Follow up more than one problem-solving alternative.
- Build up evaluation criteria to be used in iterative processes of error correction.
- Look for similar cases to be used as models for the current ones.

The practical consequences of these considerations are easy to describe. Although it is known that high specific and concrete outcome goals lead to high performance, it is not always possible to set such goals, especially when the outcome of work is not quite clear at the beginning. In such a case the chances for success will increase through setting learning goals such as the following:

- Discover a number of strategies to learn how to perform the task.
- Develop systematic testing of task-relevant strategies.
- Identify evaluation criteria.
- Use information from similar cases.

After having learned about strategies to solve the task, it could be helpful to again relate goals to outcomes.

6. MANAGING SUCCESSFUL GOAL-ORIENTED WORK BEHAVIOR: THE PARTICIPATORY PRODUCTIVITY MANAGEMENT

During the second half of the 20th century, many management systems that focused on goal setting were presented and used. The management by objectives system has been used successfully for five decades. Goal-oriented management systems help leaders to coordinate the goal-setting process continuously. The Balanced Scorecard (BSC) is well known for its ability to support successful management in this respect. Another powerful instrument with a more solid empirical foundation is the Participative Productivity Management (PPM), originally developed by Pritchard in 1990 as the Productivity Measurement and Enhancement System (ProMES). PPM is different from BSC in its very elaborate bottom-up approach that makes sure that workers accept the measurement and feedback system used to prepare adequate goal setting. So, it effectively helps to avoid goal conflicts between external or task-driven goals and personal goals set by the workers themselves.

In the framework of this management system, work group members develop their own measurement system coordinated with all responsible actors in the company, enabling them to participate as well-informed partners in the goal-setting process. In weekly or monthly systematic feedback sessions, group members and their leaders come together to discuss feedback delivered by the measurement

system that they had developed themselves. On the level of various performance indicators (e.g., amount of items produced, quality of work, customer satisfaction, scheduling effectiveness), group members use this system to describe the discrepancies between their previously set goals and their actual outcomes for each indicator. The system also helps the workers to weigh each indicator score with respect to its importance for the organization's productivity. On this basis, they discuss the reasons why they succeeded or failed to reach their goals. As a result of this discussion, they set new goals for the next measurement period. These procedures keep them in a continuous improvement process. The task of their leaders becomes more and more one of facilitators of a self-controlled productivity process. Meta-analyses show that using this management system has an enormous effect on productivity.

7. COMPETENCIES IN MANAGING GOAL-DIRECTED BEHAVIOR

Empirical data show that goal setting can be used as a technique to manage people effectively toward the goals of an organization. As was shown in the preceding section, the goal-setting process will profit from the support of a goal-oriented management system. For example, PPM offers the structural systemic background for successful goal setting. It helps leaders to do the following:

- Set goals participatively so that acceptance of organizational goals can be expected.
- Delegate responsibilities to work group members.
- Clarify goal–outcome relations.
- Define clear criteria for success and failure.

But PPM does not replace leaders. The tasks of leaders will change, but leadership will still be necessary. Leaders using a management technique such as PPM should have some special competencies to support workers' goal setting and the effect of goals during various phases of the action process. During the first phase of the action process, when the goals are developed, special competencies are required to identify the personal goals (and motives) of workers and to take them into account while discussing the necessary goals for the next work and feedback period defined in the PPM measurement system (Fig. 6). Another competency is required to structure complex tasks and identify subtasks with corresponding subgoals. In the case of evaluating negative outcomes, leaders need some competencies as well. Dealing with failure can be dangerous for leaders because it is necessary to feed back failure without trying to make things seem better as they are. The danger is that feeding back such information about failure can be seen as an indicator of low ability by failing persons. This might be a reason for lowering

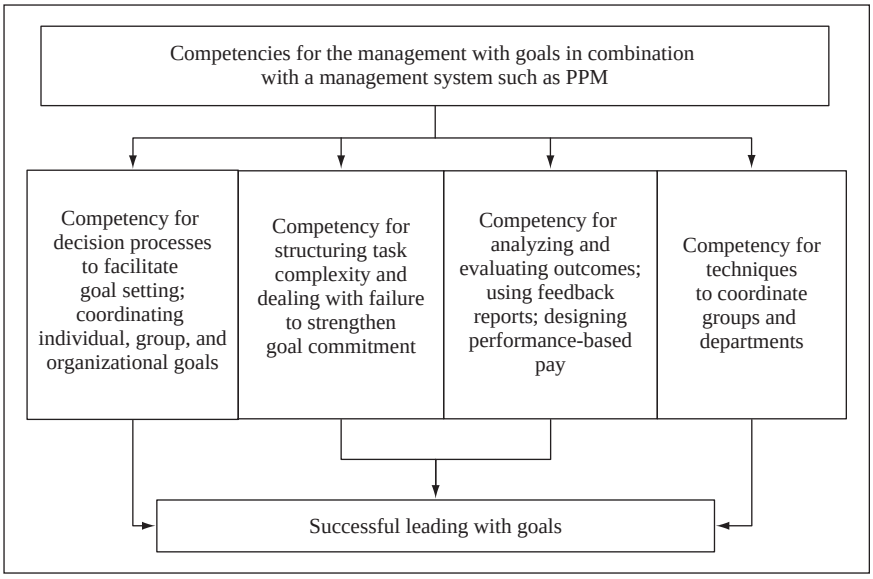

Competencies for the management with goals in combination with a management system such as PPM
Competency for decision processes to facilitate goal setting; coordinating individual, group, and organizational goals
Competency for structuring task complexity and dealing with failure to strengthen goal commitment
Competency for analyzing and evaluating outcomes; using feedback reports; designing performance-based pay
Competency for techniques to coordinate groups and departments
Successful leading with goals

FIGURE 6 Competencies needed or leaders using goal setting in combination with a goal-oriented management system.

one's self-esteem, with further effects on goal setting. Goals will be set lower, or high external goals will not be accepted, due to fears of failing again. Although work motivation can be decreased by failure feedback, it does not make sense to hold back such information. Instead, leaders should be able to talk about failure as an opportunity to learn about new strategies to avoid failure and, as such, to increase their self-efficacy. These competencies help to strengthen goal commitment. Another set of competencies for successful leadership involves coordinating persons in groups as well as between groups and departments.

8. AN INTEGRATIVE MODEL FOR THE EFFECTS OF MOTIVES AND GOALS

Motives and goals work together to control the intensity, direction, and persistence of behavior at work. Figure 7 summarizes the effects of motives and goals during all phases of the action process. Motives are defined as classes of goals that determine, in interaction with motivating potentials of the work situation, the kinds of goals people use to control their actions. The way in which they do this is described by expectancy–value theories. One example of an empirically sound specific theory of this kind is achievement motivation. It describes a motive (achievement motive) and a motivating potential (task difficulty) in explaining the development of individual goals in achievement situations. The translation of a goal into action and performance is the topic of goal-setting theory. Goal characteristics and mechanisms and moderators of the translating process are the determinants of outcomes. During this action phase, motives play a role as well. They affect the degree of goal commitment persons show while performing. Even the evaluative phase of an outcome motive can be seen at work. Motives influence the way in which causal attributions are used to explain the outcome. Especially interesting is the impact of motives on the way in which people deal with failure.

Failure delivers two kinds of information. First, it informs about the discrepancy between an outcome and the goal. Second, it says something about a person's efficiency. In case of failure, this second kind of information is dangerous for one's self-esteem. It tells a person that he or she is not able to succeed at this kind of task, especially if the failure repeats. Persons with a high achievement motive are better equipped to deal with failure because they tend to increase motivation, stay with their goals, and finally reach them. The integrated model helps to understand how goals are established, how they affect ongoing action, and how they influence the evaluation of outcomes of actions. It

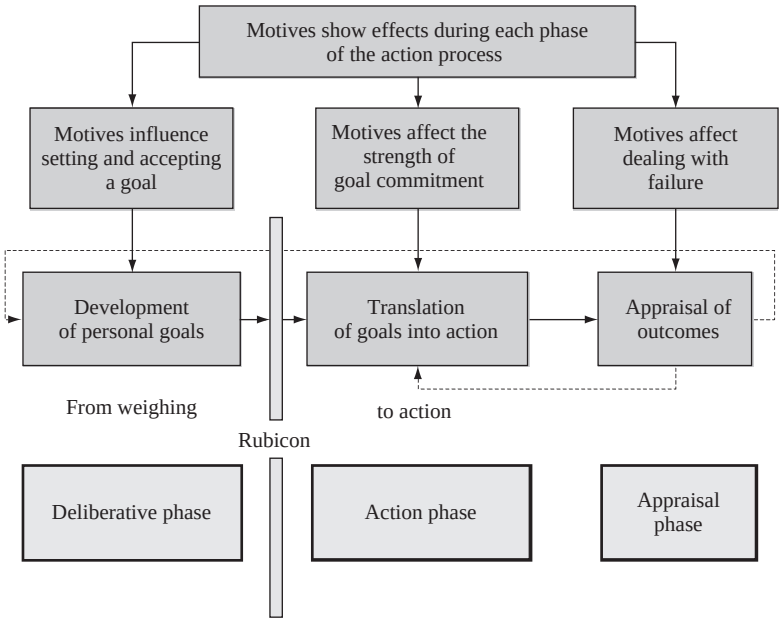

Motives show effects during each phase of the action process
Motives influence setting and accepting a goal
Motives affect the strength of goal commitment
Motives affect dealing with failure
Development of personal goals
Translation of goals into action
Appraisal of outcomes
From weighing
to action
Rubicon
Deliberative phase
Action phase
Appraisal phase

FIGURE 7 The role of motives in each phase of an action process.

summarizes a fundamental process that can be differentiated during all of its phases to describe the many facets of goal-directed behavior. Most of the concepts in the model have been tested empirically and have proven to be valid. On such an empirical ground, the model allows one to develop psychologically sound techniques to be used in applied psychology for increasing productivity and for adapting tasks and organizational contexts to the structure and functions of human resources at work.

See Also the Following Articles

Compensation ■ Cooperation at Work ■ Goal Setting and Achievement Motivation in Sport ■ Intrinsic Motivation and Self-Determination ■ Motivational Taxonomies ■ Work Motivation ■ Work Role, Values Sought in the

Further Reading

Atkinson, J. W. (1974). Strength of motivation and efficiency of performance. In J. W. Atkinson, & J. O. Raynor (Eds.), *Motivation and achievement* (pp. 193–218). Washington, DC: Winston.

Austin, J. T., & Vancouver, J. B. (1996). Goal constructs in psychology: Structure, process, and content. *Psychological Bulletin*, *120*, 338–375.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Carver, C. S., & Scheier, M. F. (1998). *On the self-regulation of behavior*. Cambridge, UK: Cambridge University Press.

Earley, P. C., Connolly, T., & Ekegren, G. (1989). Goals, strategy development, and task performance: Some limits on the efficacy of goal setting. *Journal of Applied Psychology*, *74*, 24–33.

Earley, P. C., & Shalley, C. E. (1991). New perspectives on goals and performance: Merging motivation and cognition. In G. R. Ferris, & K. M. Rowland (Eds.), *Research in personnel and human resources management* (Vol. 9, pp. 121–157). Greenwich, CT: JAI.

Hacker, W. (1998). *Allgemeine Arbeitspsychologie: Psychische Regulation von Arbeitstätigkeiten*. Bern, Germany: Hans Huber.

Hacker, W. (2004). Wie verfolgt man Ziele, die man noch nicht kennt? Ergebnisorientierte opportunistische Tätigkeitsregulation. In J. Wegge, & K-H. Schmidt (Eds.), *Förderung von Arbeitsmotivation und Gesundheit in Organisationen* (pp. 27–37). Göttingen, Germany: Hogrefe.

Hackman, J. R., & Oldham, G. R. (1976). Motivation through the design of work: Test of a theory. *Organizational Behavior and Human Performance*, *16*, 250–279.

Heckhausen, H. (1989). *Motivation und Handeln*. Berlin, Germany: Springer.

Ilgen, D. R., & Davies, C. A. (2000). Bearing bad news: Reactions to negative performance feedback. *Applied Psychology*, *49*, 550–565.

Kanfer, R., & Ackerman, P. (1989). Motivation and cognitive abilities: An integrative aptitude–treatment interaction approach to skill acquisition. *Journal of Applied Psychology*, *74*, 657–690.

Kleinbeck, U., Wegge, J., & Schmidt, K-H. (2001). Work motivation and performance in groups. In M. Erez, U. Kleinbeck, & H. Thierry (Eds.), *Work motivation in the context of a globalizing economy* (pp. 181–191). Mahwah, NJ: Lawrence Erlbaum.

Kuhl, J. (2001). *Motivation und Persönlichkeit*. Göttingen, Germany: Hogrefe.

Latham, G. P., & Frayne, C. A. (1989). Self-management training for increased job attendance: A follow-up and a replication. *Journal of Applied Psychology*, *74*, 411–416.

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting and task performance*. Englewood Cliffs, NJ: Prentice Hall.

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, *57*, 705–717.

McClelland, D. C. (1987). *Human motivation*. New York: Cambridge University Press.

Pinder, C. C. (1998). *Work motivation in organizational behavior*. Upper Saddle River, NJ: Prentice Hall.

Pritchard, R. D. (1990). *Measuring and improving organizational productivity: A practical guide*. New York: Praeger.

Pritchard, R. D., Holling, H., Lammers, F., & Clark, B. D. (Eds.). (2002). *Improving organizational performance with the productivity measurement and enhancement system: An international collaboration*. New York: Nova.

Rodgers, R., & Hunter, J. E. (1991). Impact of management by objectives on organizational productivity. *Journal of Applied Psychology*, *76*, 322–336.

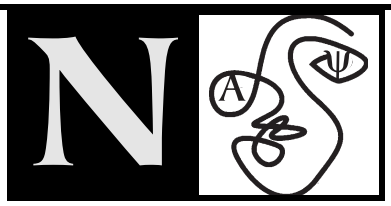
Schneider, K., & Schmalt, H-D. (2000). *Motivation*. Stuttgart, Germany: Kohlhammer.

Seijts, G. H., & Latham, G. P. (2001). The effect of distal learning, outcome, and proximal goals on a moderately complex task. *Journal of Organizational Behavior*, *22*, 291–307.

Vroom, V. H. (1964). *Work and motivation*. New York: John Wiley.

Weiner, B., Heckhausen, H., Meyer, W. U., & Cook, R. E. (1972). Causal ascriptions and achievement motivation: A conceptual analysis of effort and reanalysis of locus of control. *Journal of Personality and Social Psychology*, *21*, 239–248.

Winters, D., & Latham, G. P. (1996). The effect of learning versus outcome goals on a simple versus a complex task. *Group and Organization Management*, *21*, 235–250.



Neuropsychological Assessment in Schools

Daniel C. Miller

Texas Woman's University, Denton, Texas, USA

1. School-Based Problems That Can Be Addressed by Neuropsychological Assessment
2. State of the Art of Neuropsychological Assessment in Schools
3. Conceptual Framework for School Neuropsychological Assessment
4. Summary
Further Reading

GLOSSARY

pediatric neuropsychologist Uses standardized tests and observes behavior to define a child's pattern of cognitive development. The pediatric neuropsychologist uses knowledge of brain development, brain organization, and the effects of various forms of brain injury on development to guide this assessment and interpret results.

school neuropsychologist Integrates neuropsychological and educational principles in assessments and interventions with infants, children, and adolescents to facilitate learning and behavior within the school and family environments. School neuropsychologists also play an important role in curriculum development, classroom design, and the integration of differentiated instruction that is based on brain–behavior principles in order to provide an optimal learning environment for every child.

specific learning disabilities Disorder in one or more of the basic psychological processes involved in understanding or using spoken or written language that may manifest itself in an imperfect ability to listen, think, speak, read, write, spell, or do mathematical calculations.

traumatic brain injury An acquired injury to the brain caused by an external physical force—resulting in total or partial functional disability, psychosocial impairment, or both—that adversely affects a child's educational performance. The term applies to open or closed head injuries resulting in impairments in one or more areas, such as cognition; language; memory; attention; reasoning; abstract thinking; judgment; problem solving; sensory, perceptual, and motor abilities; psychosocial behavior; physical functions; information processing; and speech. The term does not apply to brain injuries that are congenital or degenerative or to brain injuries induced by birth trauma.

Increasingly, school-aged children and adolescents have cognitive processing deficits (e.g., memory, learning, and attention) or behavioral/emotional deficits (e.g., autism, mood disorders, anxiety, and conduct problems) that are a result of known or suspected neurological conditions. Educational personnel, including school psychologists, are not always prepared to assess and intervene with these neurologically impaired populations; however, they can be trained. Traditionally, psychoeducational assessments measure estimates of overall intelligence and academic achievement but do not assess the underlying cognitive processes and sensory–motor

©2004 Elsevier Inc.
All rights reserved.

functions associated with learning and behavior. Neuropsychological assessments are designed to assess discrete cognitive processes (e.g., visual short-term memory, auditory short-term memory, and sustained attention) that help facilitate or inhibit the learning process. Ideally, a neuropsychological assessment of a school-aged child should link the assessment results with instructional and behavioral recommendations or interventions that improve the child's academic and/or social–emotional/behavioral functioning. This article presents several examples of children who would benefit from a neuropsychological evaluation, discusses several reasons why there is growing interest in integrating neuropsychological principles into psychoeducational assessment and intervention, discusses the current state of the art of school neuropsychological assessment, and presents a conceptual framework for school neuropsychological evaluations.

1. SCHOOL-BASED PROBLEMS THAT CAN BE ADDRESSED BY NEUROPSYCHOLOGICAL ASSESSMENT

1.1. Case Examples

Increasingly, children and adolescents are coming to school with known or suspected neurological conditions. Two examples are presented here that illustrate the types of cases being identified by educators and assessment specialists within the schools.

First, Mary is 17 years old and a junior in high school. While on a date with her boyfriend, a drunk driver struck the car in which they were riding head on. Her boyfriend escaped the accident with minor injuries, but Mary was not wearing her seat belt. She was thrown through the windshield. Mary suffered a severe head injury and was hospitalized. Several weeks later, after her medical needs had become stabilized, she was released from the hospital. Her medical insurance did not cover any cognitive rehabilitation so she was sent home. When Mary returned to school a month after the accident, her physical wounds had healed and she looked as if she had healed completely. She returned to her 11th-grade regular education classes. Prior to the accident, Mary was an "A" and "B" student. When Mary returned to school after the accident, her grades declined dramatically, she complained about having trouble focusing attention and concentrating for long periods of time, she could not remember things, and she seemed to tire easily.

Second, Sam is in the fifth grade. He was diagnosed 3 years earlier with a specific learning disability in the areas of reading comprehension and written expression. Sam has average cognitive capabilities, but he has severe problems with understanding what he has read and putting his thoughts down on paper. Sam has made very little academic progress during the past 3 years. His teachers express frustration that they cannot seem to "find the right buttons to push" to help Sam make academic progress.

Sam and Mary represent two examples of children who have disabilities defined in federal legislation that are of particular interest to neuropsychologists: specific learning disability (SLD) and traumatic brain injury (TBI). In 1975, the U.S. Congress passed landmark legislation with the Education of All Handicapped Children Act (P.L. 94-142), designed to provide free and appropriate educational services to children. P.L. 94-142 was reauthorized in 1990 and became known as the Individuals with Disabilities Act (IDEA). IDEA was reauthorized in 1997 and is under revision with an expected adoption date in 2004.

1.2. Specific Learning Disabilities

SLD was included in the 1975 P.L. 94-142 as a student disability. SLD is a generic term that refers to a heterogeneous group of disorders due to identifiable or inferred central nervous system dysfunction. Such disorders may be manifested by delays in development and/or difficulties in one or more of the following areas: reading accuracy, reading comprehension, mathematical calculations, mathematical reasoning, written expression, listening comprehension, and oral expression. Since the inception of the concept of a SLD, educators have struggled with how to identify children who have these processing disorders caused by some type of central nervous system dysfunction. Commonly used methods of SLD identification, such as aptitude/achievement discrepancies, have failed to address the underlying brain processing disorder(s).

The revision of IDEA under consideration encourages states to deemphasize the use of discrepancy models in the identification of SLD and move toward a three-tiered intervention model. Tier 1 operates on the principles that high-quality instructional and behavioral supports are required for all students in general education. Tier 2 would require targeted, intensive prevention or remediation services for students whose performance and

rate of progress lag behind the norm for their grade and educational setting. Tier 3 would require a comprehensive evaluation by a multidisciplinary team to determine eligibility for special education and related services. Within the proposed third tier of the intervention model, neuropsychological assessments could provide valuable information about the child's cognitive strengths and weaknesses and the links to prescriptive interventions.

1.3. Traumatic Brain Injury

In the 1990 revision of IDEA, TBI was included for the first time as an identifiable student disability. TBI is an injury to the brain caused by the head being hit or shaken violently. When a head injury occurs in a child or adolescent, the brain has not fully developed. As a result, recovery from loss of function due to a head injury is often uneven and unpredictable. School-aged children who have sustained a head injury require organized, cooperative, and closely monitored special education services. As recently as 10 years ago, a child with a head injury would typically move from an acute care hospital setting, in which his or her physical needs were addressed, to an intermediate rehabilitation setting for a few months, in which cognitive rehabilitation took place. Today, it is typical for a child to be released from a hospital as soon as his or her medical condition is stabilized and to return to his or her regular classroom soon thereafter, without any cognitive rehabilitation. During the past 15 years, managed care within the insurance industry has lead to a reduction of cognitive rehabilitation services offered to TBI victims. Educators have been faced with how to educate children with increasingly complicated learning and emotional needs.

Neuropsychological assessment can play an important role with TBI students by determining the mental and behavioral effects of the brain injury, identifying the functional assets and deficits for instructional planning, and establishing baseline data to evaluate subsequent change. Because of the rapid changes in recovery of function, it is often necessary to meet frequently as an individualized education plan team to review the TBI student's progress and to adjust the educational goals. Educational personnel, including school psychologists, speech and language pathologists, occupational therapists, physical therapists, adaptive physical education specialists, and regular and special education teachers, must all work collaboratively with the family to meet the educational needs of the student.

1.4. Other Disabilities That May Warrant a Neuropsychological Evaluation

Table I presents a list of other suspected or known disabilities or neurological impairments that would benefit from neuropsychological assessments. In the 1990 and subsequent IDEA definitions of TBI, only injuries to the brain as a result of trauma are classified as TBI. Acquired brain injuries as a result of infection, disease, chemotherapy, or high fever are not considered under the current definition as TBI. Children with acquired brain injuries are typically identified as "other health impaired" if there are significant academic and/or behavioral concerns that need to be addressed.

Research has shown that children born with neurodevelopmental risk factors, such as low birth weight, prenatal exposure to drugs or alcohol or both, or exposure to environmental toxins (e.g., lead), are likely to present with neurocognitive deficits (e.g., memory, attention, processing speed, and language skills). Children with medical conditions such as epilepsy or children undergoing chemotherapy due to cancer may also be at risk for neurocognitive deficits associated with their illnesses or treatments.

TABLE I
Suspected Handicaps and Neurological Conditions for Which Neuropsychological Assessment May Be Beneficial

Children and adolescents who are having academic or behavioral difficulties due to
Severe learning disabilities that have not responded to traditional intervention techniques
Processing discrepancies on psychoeducational measures (e.g., memory deficits, language processing deficits, attentional processing deficits, and processing speed deficits)
A documented rapid decline in academic achievement that cannot be explained by social, environmental, or psychological causes
Congenital or acquired brain damage
Neuromuscular diseases such as cerebral palsy
Chronic epilepsy
Neurodevelopmental risk factors such as low birth weight, prenatal exposure to drugs or alcohol or both, or exposure to environmental toxins (e.g., lead)
A treated or suspected brain tumor
An infection of the central nervous system
An unusually large scatter among test scores that cannot be explained

Unfortunately, many children who have known or suspected neurological conditions are misclassified [often as learning disabled (LD) or emotionally disturbed (ED)], which leads to planning that does not fully address their educational needs. For example, an undiagnosed TBI student can develop secondary emotional and behavioral problems due to the head injury and the associated cognitive deficits. Labeling this TBI student as emotionally disturbed and focusing the majority of the educational efforts on regulating the student's behavior will not address the underlying cognitive processing problems.

Traditional psychoeducational assessments conducted with school-aged children do not fully address the types of cognitive processing deficits associated with children having known or suspected neurological conditions. Neuropsychological assessments measure discrete cognitive processes (e.g., visual short-term memory) that help facilitate or inhibit the learning process. The results of a neuropsychological assessment can provide a profile of a child's functional strengths and weaknesses that will facilitate instructional planning.

2. STATE OF THE ART OF NEUROPSYCHOLOGICAL ASSESSMENT IN SCHOOLS

2.1. Neuropsychological Assessment in the Past and Present

From the 1940s to the 1970s, one of the main functions of a neuropsychological evaluation was to identify the source of a possible lesion or area of dysfunction within the brain. With the advent of modern neuroimaging (e.g., computerized axial tomography, magnetic resonance imaging, functional magnetic resonance imaging, and positron emission tomography), these specialized techniques are now used to directly assess the localization of lesions or tumors to millimeter precision. The goal of neuropsychological assessment should be to identify the functional strengths and weaknesses of the examinee and to translate those data into evidenced-based intervention strategies rather than to try to localize possible lesion sites within the brain. A lengthy neuropsychological report to the schools with the conclusion that "Johnny has a right parietal lobe lesion" will not facilitate educational planning. The neuropsychological report must summarize implementable prescriptive recommendations that address the student's functional strengths and weaknesses for parents and teachers.

Evaluators who wanted to administer neuropsychological measures to children prior to the 1990s were very limited in their assessment choices. The fixed batteries that were standardized 20–40 years ago (e.g., Halstead–Reitan Neuropsychological Batteries and Luria–Nebraska Neuropsychological Battery) were the principle measures available for neuropsychologists. Unfortunately, these test batteries were not designed for children but were downward extensions of adult measures. Also, the fixed batteries of the 1950s and 1960s had limited standardization samples for children and would not stand up to the scrutiny of test development today. In the 1990s, test authors and publishers provided school-based assessment specialists with a wide variety of new tests of memory and learning designed for and standardized on large numbers of children and adolescents (Table II).

Since the late 1990s, there has been an increased number of fixed battery tests developed specifically for school-aged children (Table III). Some of the tests, such as the NEPSY, are neuropsychological batteries based on modern theory and research and designed specifically for young children. Other fixed batteries are intelligence tests that have been reformulated to take into account advances in cognitive psychology, neuropsychology, and intelligence test theories [e.g., Woodcock–Johnson Revised Tests of Cognitive Abilities (WJIII) and the Wechsler Intelligence Scale for Children—Fourth Edition]. Tests such as the Delis–Kaplan Executive Functions System have revised and restandardized traditional tests (e.g., Trail Making and Stroop) and added to their interpretative depth.

TABLE II
Tests of Memory and Learning for School-Aged Children

Test	*Age range (years)*
Wide Range Assessment of Memory and Learning (WRAML; Sheslow & Adams, 1990: WRAML-2; Adams and Sheslow, 2003)	WRAML: 5–17 WRAML2: 5–90
Test of Memory and Learning (TOMAL; Bigler and Reynolds, 1994)	5–19
Children's Memory Scale (CMS; Cohen, 1997)	5–16
Wechsler Memory Scale—Third Edition (WMS-III; Wechsler, 1997)	16–89
California Verbal List Learning—Second Edition (CVLT-II: Delis *et al.*, 2000)	5–16

TABLE III
Selected Fixed Batteries for School-Aged Children and Adolescents

Test	*Age range (years)*
Cognitive Assessment System (CAS; Naglieri & Das, 1997)	5–17
NEPSY (Korkman *et al.,* 1997)	3–12
WISC-III as a Process Instrument (WISC-III PI; Kaplan *et al.,* 1999)	6–16
Delis–Kaplan Executive Functions System (D-KEFS; Delis *et al.,* 2001)	8–89
Woodcock–Johnson III Tests of Cognitive Ability and Achievement (WJIII; Woodcock *et al.,* 2000)	2–90+
Dean–Woodcock Neuropsychological Battery (DW; Dean and Woodcock, 2002)	4–90+
Diagnostic Supplement to the WJIII Tests of Cognitive Abilities (DS-WJIII; Woodcock *et al.,* 2003)	2–90+
Wechsler Pre-School and Primary Scale of Intelligence—Third Edition (WPPSI-III; Wechsler, 2002)	2–7
Wechsler Intelligence Scale for Children—Fourth Edition (WISC-IV; Wechsler, 2003)	6–16
Kaufmann Assessment Battery for Children—Second Edition (KABC-II; Kaufman and Kaufman, 2004)	3–18

There are three major trends in neuropsychological assessment. The first is a subtle shift toward providing the evaluator qualitative information about the cognitive processes a child used to obtain a score, in addition to the score itself. For example, publishers are providing base rates for each age range within the standardization sample for common qualitative behaviors (e.g., asking for repetitions and failure to comprehend). This "process approach" is important because the qualitative behaviors often provide valuable insight into how a child's brain is processing information and assist in the diagnoses of neuropsychological conditions.

The second trend in neuropsychological assessment is ensuring the assessment is grounded in theory. Test authors and publishers are more consistently providing a theoretical foundation for their tests, which aids in test interpretation (e.g., NEPSY based on Lurian theory or the WJIII based on the Carroll–Horn–Cattell model). Finally, the third trend will be an added emphasis during the next decade on linking assessment with evidenced-based interventions. Researchers and clinicians will need to establish and verify linkages between assessment and intervention.

2.2. Who Can Complete a Neuropsychological Assessment?

The most highly trained assessment specialists in schools are school psychologists. However, the majority of school psychologists are trained at the specialist level (master's degree +30 hours) of training, which provides minimal training in the biological bases of behavior and rarely covers neuropsychological assessment techniques. School psychologists can only conduct neuropsychological assessments if they have expertise in neuropsychology. To ensure proper expertise, school psychologists must complete an organized competency-based, postgraduate training or a doctoral program in which they can specialize in school neuropsychology.

If a school district does not have a school neuropsychologist on staff, which is the norm, the district will often request a neuropsychological evaluation from a clinical neuropsychologist in private practice or working in a hospital setting. Clinical neuropsychologists are licensed psychologists who are trained at the doctoral level, typically with a degree in clinical psychology and a specialization in clinical neuropsychology. However, the majority of clinical neuropsychologists have expertise in completing adult assessments rather than school-based neuropsychological evaluations.

A pediatric neuropsychologist is a clinical neuropsychologist specifically trained to work with children. If school districts seek an outside neuropsychological assessment, it is best to contract a pediatric neuropsychologist. The downside to contracting with any specialist outside of schools is that the specialist may not be familiar with the IDEA rules and regulations that guide special education placements and may operate under a clinical model using clinical nomenclature (e.g., *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition, diagnoses).

For neuropsychological principles to be best integrated into schools, there needs to be more cross-training. Many of the new cognitive ability tests and tests of memory and learning being administered by school psychologists have strong foundations in neuropsychological theory. School psychologists and educators need to improve their knowledge base about brain–behavior relationships if they are going to

appropriately interpret these newer tests. There are more than 25,000 school psychologists who have direct access to children in schools. Training to use and interpret neuropsychological assessments is important. At a minimum, school psychologists and educators should be taught to recognize the warning signs of neuropsychological conditions and know when to refer to another specialist for assistance.

As a general rule, pediatric neuropsychologists consulting with schools should learn to integrate their neuropsychological assessments within the context of IDEA eligibility and recommend realistic school-based interventions. A school neuropsychology specialist can write an insightful report about a child's cognitive strengths and weaknesses and make a set of educational recommendations that are evidence based and practical. It is possible that the neuropsychological assessment report may simply be filed away in a child's academic folder if the classroom teachers do not understand how to implement the educational recommendations. An essential role that a school neuropsychology specialist can play is to consult with classroom teachers regarding the integration of the neuropsychological assessment results into the instructional planning.

3. CONCEPTUAL FRAMEWORK FOR SCHOOL NEUROPSYCHOLOGICAL ASSESSMENT

A neuropsychological evaluation for a school-aged child should focus on answering the referral question. A conceptual framework for a school neuropsychological evaluation is outlined in Table IV. A sensory perceptual examination across the auditory, visual, and kinesthetic modalities should be a required baseline assessment in all evaluations. If the child is not able to adequately receive and process sensory perceptual information, then performance on higher order cognitive functions could be compromised. The next major component to be assessed is attentional processing skills. Attention serves an important baseline for the rest of the higher order processing. Attention is not a unitary construct. It is recommended that a school neuropsychological evaluation assess the following elements of attention for both the visual and auditory modalities: initiating, sustaining, inhibiting, shifting, and divided.

Assessment of executive functions is also a recommended component. For a thorough evaluation of executive functions, the assessment must include cognitive

TABLE IV
Suggested Components of a School Neuropsychological Assessment[a]

- Sensory perceptual
  - Auditory
  - Visual
  - Kinesthetic/tactile
- Attention
  - Auditory (initiating, sustaining, inhibiting, shifting, and divided)
  - Visual (initiating, sustaining, inhibiting, shifting, and divided)
- Executive functions
  - Cognitive (e.g., processing speed and efficiency, initiating and shifting cognitive sets, conceptual shift, planning, organizing, self-monitoring, and verbal/nonverbal retrieval fluency)
  - Motor programming
  - Behavioral/emotional regulation [e.g., interest, motivation, initiative, disinhibition (impulsivity), affect, judgment, insight, and activity level]
- Memory
  - Auditory modality
    - Auditory, verbal
      - Explicit acquisition
        - Rote (word lists) (immediate and delayed)
        - Associative (paired learning) (immediate and delayed)
        - Discourse (words, sentences, paragraphs, and stories) (immediate and delayed)
      - Implicit acquisition
        - Incidental (passive)
        - Retrieval (recall—primacy and recency effects—and recognition)
    - Auditory, nonverbal
      - Explicit acquisition (immediate and delayed)
      - Implicit acquisition
      - Retrieval (recall—primacy and recency effects—and recognition)
  - Visual modality
    - Visual, verbal [letters, words, picture or shapes with verbal labels (rebus), etc.]
      - Acquisition (immediate and delayed)
        - Initial registration [sensory memory to short-term memory (immediate)]
        - Consolidation [short-term memory to long-term memory (delayed)]
      - Retrieval (recall and recognition)
    - Visual, spatial (e.g., Rey–Osterieth)
      - Acquisition (immediate and delayed)
      - Retrieval (recall and recognition)

Continues

Continued

- Motor memory
  - Procedural [e.g., literal apraxias (imitation) and paragraphias (spelling)]
    - Complex motor tasks (e.g., riding a bicycle and tying shoes)
  - Ideomotor and ideational apraxias
  - Writing and typing
- Language
  - Receptive (phonology, semantics, grammar/syntax, pragmatics, and listening comprehension)
  - Expressive [phonology, semantics, grammar/syntax, pragmatics, and discourse (e.g., articulation)]
  - Reading (phonology, semantics, grammar/syntax, pragmatics, decoding, and comprehension)
  - Written expression (not writing) (phonology–sequential processing, semantics, grammar/syntax, pragmatics, decoding, comprehension, and organization of thought and content)
- Motor
  - Fine (coordination and complex mechanical processes)
  - Gross
  - Speed
  - Strength
- Mathematics
  - Verbal (e.g., algebra)
    - Number knowledge
    - Counting
    - Computation/problem solving
    - Reasoning
  - Nonverbal (e.g., geometry)
    - Math notation [e.g., knowledge of math signs (+, −, ×, /)]
    - Computation
    - Reasoning in the time–space continuum
- Social–emotional
  - Personality characteristics
  - Temperament
  - Psychopathology

[a]Adapted from a proposed model by Miller *et al.* (manuscript in preparation).

integrity, behavioral regulation, and motor programming components. Memory assessment is next in the conceptual framework. Memory is multifaceted as well. A comprehensive evaluation must include assessment of auditory and verbal memory and the associated subcomponents (immediate, delayed, recognition, recall, etc.).

Language assessment is next in the conceptual framework. Language assessment includes measures of receptive, expressive, reading, and written expression. An evaluation of motor functions, if warranted, would include fine, gross, speed, and strength components. Mathematical skills are conceptualized as a separate area that includes verbal components (number knowledge, counting, computation, and reasoning) and nonverbal components (math notation, reasoning in time and space, and computation). Finally, a comprehensive school neuropsychological assessment should include assessment of the student's social–emotional functioning (e.g., personality characteristics, temperament, and psychopathology). This conceptual framework serves as a starting point when considering a comprehensive school neuropsychological evaluation. The component parts that are selected should vary from student to student based on the referral questions to be answered and the presenting behavioral and/or learning problems.

4. SUMMARY

Children and youth who have known or suspected neurological conditions are increasingly being identified in schools. Traditional psychoeducational measures do not fully address the underlying neurocognitive deficits associated with SLD, TBI, acquired brain injury, epilepsy, etc. However, neuropsychological assessments measure discrete neurocognitive functions (e.g., visual short-term memory, processing speed, and sustained attention) that provide valuable insight into the neurocognitive strengths and weaknesses of the child and guide prescriptive interventions to improve instruction or provide positive behavioral supports.

Neuropsychological assessments for school-aged children have changed dramatically in the past 60 years. The purpose of a neuropsychological evaluation has evolved from identifying potential brain lesions to identifying functional profiles of strengths and weaknesses that link to prescriptive interventions. There has been a tremendous increase in the number and quality of neuropsychological assessment measures available for use with school-aged children. The strengths of these measures are their theoretical foundations, reliance on empirical studies for construct development, broad standardization samples for children, inclusion of qualitative and quantitative data, and good psychometric properties.

Preferably, neuropsychological assessments should be completed by a pediatric neuropsychologist or a school neuropsychologist. There is an urgent need for more cross-training between specialists in neuropsychology and educators to ensure that the recommendations that are made in a neuropsychological assessment report can

be implemented by classroom teachers. A conceptual framework for a school-based neuropsychological evaluation was also presented in this article. The selected component parts should be based on the referral questions and the presenting behavioral and/or learning problems.

The application of neuropsychological assessments in schools is an emerging area within the broader field of psychology and school psychology. For neuropsychological principles to be fully integrated within schools, brain-based research needs to be more fully translated into practice. More specialists in neuropsychology need to be trained to be readily available to educators for consultation, assessment, and prescriptive intervention.

See Also the Following Articles

Clinical Assessment ■ Diagnostic and Statistical Manual of Mental Disorders ■ Emotional and Behavioral Problems, Students with ■ Psychological Assessment, Standards and Guidelines for ■ Psychometric Tests ■ Psychoneuroimmunology ■ Psychophysiological Assessment

Further Reading

Baron, I. S. (2003). *Neuropsychological evaluation of the child*. New York: Oxford University Press.

Feifer, S. G., & DeFina, P. A. (2000). *The neuropsychology of reading disorders*. Middletown, MD: School Neuropsych Press.

Feifer, S. G., & DeFina, P. A. (2002). *The neuropsychology of written language disorders*. Middletown, MD: School Neuropsych Press.

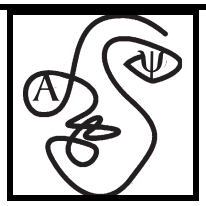
Hebben, N., & Milberg, W. (2002). *Essentials of neuropsychological assessment*. New York: Wiley.

Levine, M. D., & Reed, M. (1998). *Developmental variation and learning disorders* (2nd ed.). Cambridge. MA: Educators Publishing Service.

Reynolds, C. R., & Fletcher-Janzen, E. (Eds.). (1997). *Handbook of clinical child neuropsychology* (2nd ed.). New York: Plenum Press.

Semrud-Clikeman, M. (2001). *Traumatic brain injury in children and adolescents: Assessment and intervention*. New York: Guilford Press.

Teeter, P. A., & Semrud-Clikeman, M. (1997). *Child neuropsychology: Assessment and interventions for neurodevelopmental disorders*. Needham Heights, MA: Allyn & Bacon.



Nicotine Addiction

Neil E. Grunberg and Jennifer M. Phillips

Uniformed Services University of the Health Sciences, Bethesda, Maryland, USA

1. Introduction
2. Nicotine Overview
3. Role of Nicotine in Tobacco Use
4. Nicotine as a Treatment for Tobacco Cessation
5. Directions in Nicotine Addiction Research and Application
Further Reading

GLOSSARY

addiction A behavior pattern of drug use characterized by overwhelming involvement with the use of the drug, the securing of its supply, and a high tendency to relapse after withdrawal.

classical conditioning Learning in which the conditioned stimulus is paired with and precedes the unconditioned stimulus until the conditioned stimulus alone is sufficient to elicit the response.

dependence The highly controlled or compulsive use of a psychoactive drug that is reinforced by the effects of that drug and includes the occurrence of tolerance and withdrawal.

nicotine A chemical alkaloid ($C_{10}H_{14}N_2$) that is naturally present in tobacco.

operant conditioning Learning in which the desired behavior or increasingly closer approximations to it are followed by a rewarding or reinforcing stimulus.

tolerance The need for markedly increased amounts of a substance to achieve the desired effect or a markedly diminished effect of a drug with continued use of the same amount of the substance.

withdrawal The syndrome of often painful or uncomfortable physiological and psychological effects that result when the use of a drug is discontinued.

Tobacco use and cigarette smoking are widespread behaviors that involve the self-administration of the addictive drug nicotine. These behaviors directly contribute to premature morbidity and mortality from a variety of diseases, including cardiovascular diseases, cancers, and respiratory diseases. A great deal of information has accumulated relevant to the reasons people use tobacco products and the role of nicotine addiction. In addition, nicotine replacement therapies are used in conjunction with behavioral and cognitive therapies to help people abstain from tobacco use.

1. INTRODUCTION

Tobacco use and cigarette smoking are worldwide health problems. It is estimated that more than 1 billion adults, or one-third of the global adult population, smoke cigarettes. In the United States, approximately 46.5 million adults (23.3% of the adult population) are cigarette smokers and 7.6 million adults are smokeless tobacco users. Cigarette smoking is an especially difficult problem in the adolescent population. Although smoking among high school students is beginning to decline, more than 2000 people younger than age 18 begin smoking every day in the United States. In addition to

©2004 Elsevier Inc.
All rights reserved.

age, factors such as gender, race, and socioeconomic status affect the likelihood that an individual will use cigarettes or other tobacco-containing products.

The high prevalence of tobacco use results in significant costs to society. The medical consequences of smoking and tobacco use are extensive and result in 430,000 premature deaths per year in the United States and more than 5 million premature deaths per year worldwide, making tobacco use the leading preventable cause of death. Smoking is significantly related to an increased risk for stroke, heart attacks, vascular diseases, and aneurysms. It is estimated that approximately one-fifth of all heart disease deaths are smoking related. Tobacco use accounts for one-third of all cancers and has been associated with approximately 90% of all lung cancer cases, the leading source of cancer-related deaths in both men and women. In addition, smoking causes several other types of cancer and pulmonary diseases, and smoking increases overall morbidity rates. Smoking and tobacco use in the United States cost approximately $80 billion per year in health care costs and $138 billion per year when factors such as smoking-related fires, medical care for low-birth-weight infants born to smokers, and the medical effects of second-hand smoke are taken into account.

Psychological and biological processes contribute to the initiation, maintenance, and cessation of tobacco use. People begin to use tobacco as a result of peer pressure, role modeling, responses to advertising, and other social and psychological influences that expose them to the addictive drug nicotine, which is responsible for the development of tobacco dependence. Self-administration of nicotine via tobacco becomes interwoven with environmental stimuli. Principles of learning, including classical conditioning, operant conditioning, and paired associations, act in concert with pharmacological actions of nicotine to positively and negatively reinforce this self-administration. Treatment of tobacco use, therefore, involves pharmacological and psychological interventions that work together.

2. NICOTINE OVERVIEW

Nicotine is one of the 4000 chemical components of tobacco smoke and is also present in smokeless tobacco. It is an alkaloid that is derived from the leaves of several plants, most commonly tobacco. Nicotine's molecular formula is $C_{10}H_{14}N_2$, and structurally it is composed of a pyridine and a pyrrolidine ring. The fact that nicotine is a relatively small molecule and composed of elements that are compatible with biological systems contributes to its distribution and actions in the body. The most common route of administration of nicotine is by smoking cured and processed tobacco leaves in the forms of cigarettes, cigars, or pipe tobacco. The pH of cigarette smoke is slightly acidic (5.5–6.5) and nicotine's dissociation constant (*p*Ka)—the pH at which the chemical is in its ionized versus deionized form—is 8.0. As a result, nicotine in cigarette smoke does not readily diffuse across the mucous membranes of the mouth or throat. The smoke must instead be inhaled into the lungs to be absorbed across the membranes of the pulmonary alveoli. Nicotine passes into the pulmonary vascular circuit and is transported directly to the brain within 5–10 seconds of inhalation. The nicotine acts in the brain and is carried in the bloodstream throughout the body, metabolized in the liver, passed to the kidneys, and then eliminated from the body.

Psychological and behavioral variables (including stress, diet, and activity) can alter the distribution of nicotine and its rate of elimination from the body, which helps explain why people smoke more in certain situations (e.g., when under stress). The pH of chewing tobacco, snuff, and cigar and pipe smoke differs from that of cigarette smoke; as a result, nicotine is absorbed across mucosal membranes from these noncigarette tobacco products. When nicotine is absorbed across membranes in the mouth, rather than in the lungs, it is slower to act on the brain.

In the brain, nicotine's effects are mediated by several subtypes of nicotinic acetyl cholinergic receptors. These receptors are pentameric (five-unit) peptides that are found in the brain and central nervous system. The stimulation of these receptors by nicotine results in cholinergic activation and an increased release of dopamine and norepinephrine. In addition, nicotine affects the release and availability of endogenous opioid peptides, serotonin, and a variety of other biochemicals. The effects of nicotine on dopamine and endogenous opioid peptides are particularly relevant to nicotine's rewarding effects and play a major role in the biological mechanisms underlying nicotine addiction. It is well established that dopamine is involved in reward pathways that reinforce the self-administration of drugs of addiction, including nicotine, cocaine, and opiates. In addition, the endogenous opioid peptides are the naturally occurring biochemicals that mimic effects of the opiates, including morphine and heroin.

3. ROLE OF NICOTINE IN TOBACCO USE

To understand the role of nicotine addiction in tobacco use and cigarette smoking, it is important to recognize and define the classical triad of drug addiction: dependence, tolerance, and withdrawal. The terms drug addiction and drug dependence are often used interchangeably. According to the 1988 U.S. Surgeon General's report on nicotine addiction, drug dependence refers to the highly controlled or compulsive use of a psychoactive drug that is reinforced by the effects of that drug. The American Psychiatric Association uses a similar definition for drug dependence but also includes the occurrence of tolerance and withdrawal as necessary to establish drug dependence. Tolerance is the need for markedly increased amounts of a substance to achieve the desired effect or a markedly diminished effect of a drug with continued use of the same amount of the substance. Withdrawal is the syndrome of often painful or uncomfortable physiological and psychological effects that result when the use of a drug is discontinued.

Nicotine fits the classic profile of an addictive drug resulting in dependence as characterized by the occurrence of controlled and compulsive use as well as tolerance and withdrawal in a variety of species of animals and humans. Nicotine has been identified as the addictive drug in tobacco that is largely responsible for development of chronic tobacco use and dependence. It is an addicting drug with multiple reinforcing actions that are relevant to the maintenance of cigarette smoking and tobacco use as well as relapse during quit attempts. Cigarette smokers report positive, rewarding effects of nicotine and negative effects following abstinence from nicotine self-administration. As a result, nicotine has positive and negative reinforcing actions. These reinforcing actions are mediated by neurochemical (including dopamine and opioid peptides), electrophysiological (including brain stimulation and muscular relaxation), and psychobiological mechanisms. Administration of nicotine to animals produces similar effects and alleviates symptoms of withdrawal. Psychological processes and principles (including classical conditioning, operant conditioning, and paired associations) contribute to the self-administration of nicotine. Nicotine's biobehavioral and psychobiological effects (including body weight control, attentional focusing, and stress management) also reinforce nicotine self-administration. In addition, nicotine has potential use as a treatment for a wide variety of medical conditions.

4. NICOTINE AS A TREATMENT FOR TOBACCO CESSATION

Although nicotine causes addiction, it does not cause most of the health risks associated with tobacco use. Nicotine replacement therapies (NRTs) are logical treatments to help people abstain from tobacco use because they help to relieve withdrawal symptoms associated with smoking cessation. In addition, the lower dosages of nicotine supplied by current NRTs do not produce the pleasurable effects associated with tobacco products and abuse potential is relatively low.

Nicotine polacrilex gum was the first pharmacological agent approved for smoking cessation treatment by the Food and Drug Administration (FDA). In 1984, it was only available by prescription. A higher dosage gum was introduced in 1992. In 1996, the FDA approved nicotine gum for over-the-counter sales. Nicotine patches that deliver nicotine transdermally were introduced in the United States in 1991. Two variations of these patches became available over-the-counter in 1996.

Today, there also are nicotine nasal sprays, inhalers, lozenges, and a variety of other nicotine-containing products, such as lollipops and water, that are marketed as NRTs. The FDA estimates that more than 1 million U.S. smokers have been successfully treated for nicotine addiction since the introduction of NRTs. The availability of over-the-counter NRTs and media campaigns to increase smoking cessation have produced an estimated 20% increase in successful cessation attempts each year. In addition, there are other nonnicotine pharmacological treatments to help people abstain from smoking. It is important to note that abstinence success is maximized when pharmacological treatments are used in conjunction with behavioral, cognitive, and psychosocial treatments.

5. DIRECTIONS IN NICOTINE ADDICTION RESEARCH AND APPLICATION

Psychologists, pharmacologists, neuroscientists, molecular biologists, epidemiologists, and other health care scientists and practitioners are actively studying nicotine addiction and tobacco use with the goal of developing additional and more effective ways to prevent and treat tobacco use. Ongoing research is examining various forms of NRT, novel nicotine analogs (i.e., drugs that

are similar to nicotine), and various pharmacological treatments that alter different mechanisms underlying nicotine's actions. In addition, psychosocial and cognitive–behavioral treatments are being coupled with pharmacotherapy to optimize tobacco cessation. Progress in understanding the mechanisms and treatment of nicotine addiction is constantly being made through innovative new research and techniques in the field. Many researchers are employing various methods of brain imaging to better understand how nicotine affects specific areas of the brain and how the sites and mechanisms of action might contribute to nicotine's addictive effects. At the molecular level, a variety of strains of genetically manipulated knockout mice provide information regarding the receptor biology associated with nicotine's effects.

In addition to pharmacological and psychological studies and strategies, immunopharmacotherapy is an active focus of inquiry for tobacco treatment. Substances are being developed and tested that can immunologically seek nicotine as it enters the body, complex with it, and prevent it from binding to its receptors. Of course, such treatment requires patient compliance, so psychological variables again become relevant and important for effective treatments.

The tailoring of treatments to people who are at the greatest risk of tobacco use is also an area of emphasis in nicotine research. People with mental health conditions, including depression, anxiety disorders, and thought disorders, are at high risk for tobacco use. In fact, approximately 50% of people with mental health conditions smoke. Therefore, it is important to determine why these high comorbidities exist and how to help these people quit smoking.

Nicotine has also been proposed as a potential treatment for disorders other than nicotine addiction. Scientists are examining the use of nicotine to treat disorders such as Alzheimer's disease, dementias, attention deficit hyperactivity disorder, skeletomuscular diseases, Parkinson's disease, obesity, anxiety, affective disorders, and schizophrenia.

See Also the Following Articles

Alcohol Dependence ■ Drug Dependence

Further Reading

Benowitz, N. L. (1999). Nicotine addiction. *Primary Care*, *26*, 611–631.

Grunberg, N. E., Brown, K. J., & Klein, L. C. (1997). Tobacco smoking. In A. Baum *et al.* (Eds.), *Cambridge handbook of psychology, health and medicine* (pp. 606–611). Cambridge University Press, Cambridge, UK.

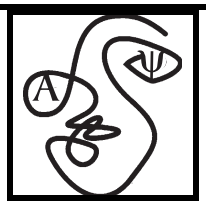
Grunberg, N. E., Faraday, M. F., & Rahman, M. A. (2001). The psychobiology of nicotine self-administration. In A. Baum *et al.* (Eds.), *Handbook of health psychology* (pp. 249–262). Mahwah, NJ: Lawrence Erlbaum.

Hays, J. T. (2000). Tobacco, nicotine and addiction. *Journal of Substance Use*, *5*, 159–164.

Henningfield, J. E. (1995). Nicotine medications for smoking cessation. *New England Journal of Medicine*, *333*, 1196–1203.

U.S. Department of Health and Human Services. (1988). *The health consequences of smoking: Nicotine Addiction. A report of the Surgeon General* (DHHS Publication No. CDC 88-8406). Washington, DC: U.S. Government Printing Office.

U.S. Department of Health and Human Services. (2001). *National Institute on Drug Abuse Research report series: Nicotine addiction* (DHHS Publication No. 01-4342). Washington, DC: U.S. Government Printing Office.



Nondiscriminatory Assessment in Schools

Samuel O. Ortiz

St. John's University, Jamaica, New York, USA

1. Introduction
2. Brief History of Nondiscriminatory Assessment in Schools
3. Issues and Definitions of Test Bias
4. Issues in Psychological Testing
5. Nondiscriminatory Assessment
6. Summary
   Further Reading

GLOSSARY

acculturation The process by which individuals from a different culture learn about and assimilate the values, beliefs, and attitudes of the mainstream culture in which they now live, as reflected by their knowledge of or participation in the new culture, including use of the dominant language. Acculturation is like any other developmental process and is largely a function of time and exposure to the mainstream culture so that individuals who have lived in the new cultural milieu longer tend to display higher levels of acculturation—that is, more knowledge of the culture and more proficiency and use of the dominant language of the culture.

assumption of comparability When tests are normed on particular populations, there is an assumption that all the individuals share similar or comparable, but not necessarily identical, background or acculturative experiences. This is the reason why 10-year-olds are compared to other 10-year-olds and not to 15-year-olds. If it can be assumed that there is comparability in the backgrounds of the individuals on whom the test was normed (namely in terms of level of acculturation and language proficiency/development), then along with other controls for variation (age, grade, socioeconomic status, region, etc.), the norms can be said to establish an appropriate base for comparing performance of individuals against the group as a whole.

bilingual Refers to any individual with some experience or exposure to at least two different languages. Being bilingual is not an either/or proposition; rather, it is a continuum that ranges from people with very low levels of proficiency in both languages to people with very high levels of proficiency in two or more languages, or balanced bilinguals. Generally, bilingual individuals will have a "dominant" or preferred language and their proficiency is not often identical; being very proficient in one language and limited in another is quite common. Nonetheless, bilingual individuals do not share and do not have comparable backgrounds to individuals who have experience primarily with only one language.

culture The idiosyncratic values, beliefs, attitudes, customs, traditions, and language that belong to a circumscribed group of people and that are transmitted as part of the collective societal knowledge and behaviors from one generation to the next. Culture thus represents a powerful environment that shapes individuals in a unique way and differently from one group to the next. Culture is often equated with or confused with race or ethnicity, but it is neither. People of varying ethnoracial backgrounds can share the same culture. Likewise, people from the same ethnoracial background can come from markedly different cultures. It is culture, not race or ethnicity, that is

©2004 Elsevier Inc.
All rights reserved.

important because it is the force that defines and shapes an individual's behavior and cognition.

psychometrics Literally means "measurement of mental abilities" and refers to the field in psychology that is devoted to establishing statistical and mathematical procedures for quantifying a wide range of human cognitive abilities. Such abilities are latent; that is, they cannot be measured directly and their existence is inferred on the basis of manifest attributes that can be directly measured.

saturation This is similar to the concept of acculturation and refers specifically to the degree to which an individual's knowledge, experiences, and development closely align with expectations for individuals of the same age in the society. When an individual has acquired this knowledge or experience (become acculturated) to the degree that would be considered developmentally appropriate, his or her level of saturation can be said to be normal or comparable to that of other individuals of the same age or grade.

test validity The degree to which a test is valid is an indication of the degree to which it measures what it purports and is designed to measure. Validity is related to reliability, or the accuracy with which a test can measure something, because a test must first be reliable (or accurate) before it can be valid. No test that is unreliable or inaccurate can be reliable.

Although the need and rationale behind nondiscriminatory assessment have been clear from the outset of the development of psychometrics, the actual practice of providing fair and equitable assessment has yet to be fully embraced by practitioners in school psychology and other major areas of psychological practice. Indeed, many of the issues related to the use of standardized tests, particularly regarding individuals from diverse linguistic and cultural backgrounds, remain poorly or incompletely understood by school psychologists. There are many reasons for this gap, including lack of training standards and models of practice; a relative dearth of practitioners and academicians with expertise in the field; confusion regarding the nature of bias or discrimination in assessment; the lack of systematic guidelines for everyday practice; and the need to develop transdisciplinary, collaborative approaches that span other professional domains, such as speech–language and education. This article provides a brief overview of these issues including a historical perspective on nondiscriminatory assessment; the consequences of inequitable evaluations; the relationship between culture, language, cognition, and education; distinguishing definitions of bias in testing; and preassessment considerations. This article also specifies steps that can be taken to conduct a comprehensive and systematic evaluation that is in keeping with best practices on nondiscriminatory assessment.

1. INTRODUCTION

In its simplest form, nondiscriminatory assessment refers to assessment that is conducted in a manner so as to ensure the validity of obtained results. Generally, the tests and procedures employed in assessment have been designed for use with a wide range of diverse individuals so that adherence to standardization and other guidelines effectively render the obtained results valid. However, the issues are not quite so straightforward. Even at the very early stages of psychometric development, psychologists were well aware of the potential problems with regard to testing diverse individuals. George Sanchez (as cited in Valdés and Figueroa, 1994), a Mexican American psychologist, noted with acute insight that "as long as tests do not at least sample in equal degree a state of saturation [assimilation of fundamental experiences and activities] that is equal for the 'norm children' and the particular bilingual child it cannot be assumed that the test is a valid one for the child" (p. 106).

The issue of validity is at the heart of nondiscriminatory assessment and the failure on the part of psychologists to fully understand and delineate this variable has led to far-ranging ramifications that may be widely recognized but are often misunderstood. Despite some attention to this issue, the effects of biased and discriminatory assessment (e.g., attribution of culture-specific behavior to various forms of psychopathology, overrepresentation of African American children in special education, and misdiagnosis of the normal process of second language acquisition as a disability) continue to be seen in virtually all aspects of psychological practice, particularly in schools. How and in what manner these variables have affected perspectives and practices in nondiscriminatory assessment will be discussed further in the following sections.

2. BRIEF HISTORY OF NONDISCRIMINATORY ASSESSMENT IN SCHOOLS

It is important to recognize that testing, evaluation, and assessment are not interchangeable terms. In the

school setting, testing represents the most narrow activity and refers to the actual use of tests to measure some trait or characteristic. Evaluation is the process of making inferences and judgments based on data collected during testing and generally takes the form of diagnosis, disability determination, and program eligibility consideration. The main focus of the activities that constitute evaluation is on "how much" a person exhibits some trait relative to same age or grade peers. Assessment is the broadest activity of all and is defined generally as the process of problem solving based on information gathered during testing, any evaluative inferences made regarding test data, as well as on other data, including observation, interviews, work samples, and review of existing records. Assessment activities ultimately lead to a corrective solution and nondiscriminatory assessment ensures that the solution is responsive to the particular individuals, elements, and systems that comprise the problematic situation.

One of the first legal battles regarding issues of potential test bias occurred in California in *Diana v. State Board of Education* (1970). The gist of the lawsuit by the plaintiffs was that the "educable mentally retarded" (EMR) placement and diagnosis were incorrect, particularly because the manner by which the latter was determined was invalid—an IQ test administered in English to children who were predominantly Spanish speakers. However, the case was settled out of court and never went to trial because the California State Board of Education agreed to adopt certain specific remedies to the satisfaction of the plaintiffs. The remedies were largely ineffective but set certain precedents regarding the testing of students with limited English proficiency, including testing in the primary language and in English, use of nonverbal IQ, reevaluation of children previously diagnosed as EMR with particular attention to nonverbal performance, development of an intelligence test appropriate for Mexican American children (never accomplished), and monitoring of the representation of EMR Hispanic students in special education throughout the state.

The remedies in the Diana case demonstrated the growing concern over linguistic issues and the potential discriminatory aspects and outcomes in assessment and education in schools. This sentiment was further reinforced by the U.S. Supreme Court in its *Lau v. Nichols* (1974) ruling that Chinese-speaking students were being discriminated against by having school instruction provided in English only—a language they could not understand—and thus were not receiving an equal educational opportunity. Although the primary impact of the Lau decision was the implementation of a system for evaluating schoolchildren's linguistic proficiency and need for equitable access to the curriculum, it also carried ramifications that extended to the assessment of students with limited English proficiency. Foremost among these implications is the appreciation of the potential bias inherent in evaluating students in a language they do not comprehend or comprehend fully as well as the possible consequences of failure to evaluate performance in the native language.

In 1975, P.L. 94-142 [known as the Education of Handicapped Children Act and reauthorized several times since then as the Individuals with Disabilities Education Act (IDEA)] contained specific wording that addressed issues in assessment in schools and whether in fact it was nondiscriminatory in nature. Several stipulations in IDEA make it clear that assessment has to account for possible areas of bias or discrimination, as evidenced by the following statement that remains in current versions: "Tests and other evaluation materials used to assess a child ... are selected and administered so as not to be discriminatory on a racial or cultural basis; and are provided and administered in the child's native language or other mode of communication" (IDEA §300.532, Evaluation Procedures). The emphasis in the law on fair and equitable assessment is quite clear, as is the notion that "tests" in particular may prove to be discriminatory if not properly used. Soon after the implementation of IDEA, the question of the fairness of IQ tests emerged prominently, as did concerns regarding the apparent negative consequences of their use, as evidenced most often by the overrepresentation of one cultural or ethnic group in special education.

The potential discriminatory aspects of psychological tests and testing in schools is exemplified in the landmark California case known as Larry P. (*Larry P. v. Riles*, 1979). In this case, it was argued by the plaintiffs that the use of IQ tests with African Americans resulted in improper special education placements largely because the tests were biased. Judge Robert Peckham, the presiding jurist, provided an initial ruling that still stands today (in California) that prohibited the use of IQ tests for the purpose of determining placement in "dead-end" special education classes. In other words, IQ tests could not be used to establish the intelligence level of African American children in cases of suspected mental retardation because it had not been demonstrated that such tests had been validated for this purpose.

A later order by Judge Peckham banning the use of IQ tests altogether with African American children was rescinded on procedural grounds. However, the original order remains intact and the California Department of Education later adopted policy in keeping with the restriction of IQ tests for any special education purposes involving African American children.

At approximately the same time as the Larry P. case, the issue of IQ test bias and overrepresentation was also being considered in *Parents in Action for Special Education* (PASE) *v. Hannon* (1980). In this case, Judge John Grady limited the question of test bias to whether IQ tests were culturally biased against African American students and even examined for himself every item and correct answer from the Wechsler Intelligence Scale for Children (WISC), WISC-Revised (WISC-R), and the Stanford–Binet IQ tests. In rendering his judicial opinion, Judge Grady found only a handful of test questions (eight on the WISC and WISC-R and only one on the Binet) that he believed were biased and subsequently ruled in favor of the defendants, citing that the plaintiffs (PASE) failed to prove the question of bias in IQ tests. The more interesting aspect of the case, however, is that Judge Grady cited that in Chicago special education placement was not based primarily or exclusively on the IQ score but rather on a rigorous and careful placement process with well-trained and knowledgeable professionals, including sufficient numbers of African American school psychologists to manage the process fairly and equitably. Judge Grady's decision, although rather tenuous on the issue of test bias, reinforced the notion that nondiscriminatory assessment is a process and includes many aspects that may create or contribute to bias—all of which must be attended to in order to ensure fairness.

Criticism and concern regarding the discriminatory aspects of assessment in schools continues almost unabated primarily because of statistics that still indicate a disproportionate representation of African Americans (and other ethnic groups) in special education programs throughout the United States. Typically, African Americans are overrepresented in the Mentally Retarded and Emotionally Disturbed categories while conspicuously underrepresented in the Specific Learning Disabled and Speech–Language Impairment categories. The U.S. Office of Civil Rights has pointed to this overrepresentation as evidence of discriminatory assessment practices (which presumably include biased testing) and has routinely used its authority to enforce compliance with procedures and practices that do not result in such inequitable outcomes. Nonetheless, overrepresentation continues to be a national problem in schools.

3. ISSUES AND DEFINITIONS OF TEST BIAS

The issue of bias in testing has revolved almost exclusively around statistical or psychometrically based definitions—that is, that bias exists as an inherent flaw within the test that results in a systematic but unintended pattern of results for a particular group of individuals. Thus, bias has been examined from various aspects of test construction, including as a function of such characteristics as (i) item content or novelty; (ii) item sequence, order, or difficulty; (iii) reliability in terms of measurement error or accuracy; (iv) stability of the underlying factorial structure; (v) examiner–examinee or atmosphere effects; and (vi) predictive ability, usually related to academic achievement or success.

Various types of studies, including correlation, multiple regression, factor analysis, prediction, and other types of analyses on test characteristics, have been conducted for decades and, by and large, evidence of bias along these lines has been rather meager. For example, items on tests with clear cultural connotations have been abandoned in favor of content that is more general and universal. Sophisticated techniques for creating scales and tests whereby the items show measurably progressive levels of difficulty (e.g., Rasch scaling) have been in use for more than a decade. Virtually all IQ tests have demonstrated more than adequate forms of reliability and their underlying factor structures have not been found to change as a function of cultural or ethnically based group differences. However, perhaps the most often cited type of research with respect to issues of bias (or lack thereof) revolves around studies that examine fairness via prediction methods. Generally, a test is considered to be fair when its use does not lead to different predictions for different groups as determined by comparison of the regression lines between a predictor test. Only when the characteristics of such regression lines differ across groups is it believed that a test may be unfair. Although this method of examining bias appears intuitive and appealing, it should be noted that even perfectly matched regression lines do not automatically imply that the construct or ability underlying the test used for prediction is in fact the construct or ability that

was actually measured. A test may well predict equally well for native English speakers as it does for limited English speakers, but the nature of exactly what ability was measured in each case can differ substantially.

Overall, these results are not surprising because the sophistication and psychometric rigor of modern-day tests virtually guarantee that bias in this form will not operate because careful attention has been paid to each of these issues in the development and construction of the test. Combined with norms that are stratified on the basis of age, grade, gender, geographical location, socioeconomic status, race/ethnicity, and other assorted variables, use and application of current assessment instruments would appear to carry little risk of bias defined as such. Indeed, when bias is viewed only as a function of the psychometric properties of a test (generally related to issues of reliability), it is unlikely to be found.

However, psychometric bias is by no means that simple a matter. Tests must not only be reliable but also be valid—that is, they must measure the very constructs they were designed to measure. By definition, a test that is valid must already be reliable, but a test that is reliable is not automatically valid. Hence, the expectation that a very reliable test will show evidence of psychometric bias related to its accuracy is essentially a straw man argument because the variables that might affect reliability have already been controlled or eliminated during test development. What has not been recognized, however, is that there are powerful variables that may not affect reliability but can indeed significantly affect validity. Inequity in assessment is most likely to occur in cases in which the obtained results no longer reflect the constructs that were purportedly measured but, rather, different constructs that were unintentionally measured. In its most simple form, this concept is illustrated in the administration of a test in English to an individual with no or an incomplete understanding of English. The test no longer provides an accurate measure of ability or achievement but, rather, reflects the individual's comprehension of and proficiency in the English language. Nondiscriminatory assessment demands recognition of the fact that variations in acculturation and language backgrounds and experiences can and do affect test validity—a historically underresearched area of test bias.

According to Scarr (1978), "intelligence tests are not tests of intelligence in some abstract, culture-free way. They are measures of the ability to function intellectually by virtue of knowledge and skills in the culture of which they are a sample" (p. 339). High-quality tests such as those currently available may indeed not be "culturally biased" per se but they remain "culturally loaded"—that is, they have embedded content that is specific and often unique to the culture that gave rise to the test. For tests developed in the United States, this would encompass content and knowledge that is known as "mainstream" or popular—information that would be acquired either incidentally through exposure and contact with the culture or formally through instruction. The process by which an individual learns the common elements of a particular culture is known as acculturation, and it is a process that advances naturally and over time, as do all developmental processes. The more frequent, simple, and common elements of the culture are learned first and the more complex ones later.

Similarly, tests that rely on an individual's ability to comprehend English can be construed as being "linguistically demanding"—that is, they require certain expected levels of language proficiency or facility with the language in order to adequately or successfully understand directions for the task or even perform the task. Learning any language, particularly a second language, is also a developmental process. Usually, the process of learning a language begins at birth, but in many cases it may begin at some point after birth, putting the acquisition and learning of that language ostensibly "behind" that of native speakers. This means that the normal expectations for language proficiency (as well as those for acculturative knowledge acquisition) for a given age or grade may not yet have been attained by those who began learning the language at a point other than from birth or who were not indoctrinated into the mainstream culture from birth. Because neither language nor acculturation are finite, there is no point at which individuals who began at points other than birth will ever "catch up" or become equivalent to their mainstream, native speaking counterparts—except in cases in which individuals were well schooled and educated in their native language first, usually up to at least age 11 or the sixth grade.

In summary, both the level of acculturation and language proficiency (generally, English language proficiency) can significantly affect the process of assessment and can alter the psychometric properties of a test with regard to its validity even though the inherent reliability of the test remains intact. Tests in particular are constructed on an assumption of comparability, which presumes a relatively similar background and experiences for the individuals on whom the test was normed. Variations in background or experiences that violate this assumption therefore render test results

effectively invalid. This is most likely to occur in cases involving culturally or linguistically diverse individuals because the test may well degenerate into one that measures level of acculturation or English language proficiency more so than the intended or expected constructs. In such situations, the tests used may not be biased, but they are not likely valid either.

Apart from the more commonly defined areas of bias, there are other definitions that are important relative to nondiscriminatory assessment, albeit they are perhaps not as widely accepted. Bias in tests is often cited as a function of the fact that a test results in (i) mean differences (generally between different cultural or ethnic groups); (ii) variable selection ratios or disproportionate classification, usually in special education; and (iii) social consequences. Indeed, mean differences and disproportionate classification formed the basis of the arguments that prompted the advancement of the historical legal challenges to the use of IQ tests in schools discussed previously. Although the result may well be pernicious, it is important to note that bias definitions of this type reflect more the application and use of tests than any inherently discriminatory property of tests. This shift in the nature and definition of bias is particularly evident with respect to the issue of social consequences, or more accurately the negative social consequences of the use of IQ tests, which generally arises out of recognition of a more general failure on the part of special education to provide positive outcomes. Frequently, use of IQ tests leads to disability classification, which then leads to placement in special education organized around the disability label. When combined with the negative findings on attempts to match instruction to cognitive or neuropsychological strengths, the efficacy of labeling children via use of IQ tests results leads to decidedly negative and discriminatory social consequences.

4. ISSUES IN PSYCHOLOGICAL TESTING

Perhaps the area that psychologists tend to focus on most when considering nondiscriminatory assessment in general is that which pertains to what is the best test to use. This may be viewed as somewhat of a search for the Holy Grail. Sattler (1992) expressed this illusion best when he stated that "probably no test can be created that will entirely eliminate the influence of learning and cultural experiences. The test content and materials, the language in which the questions are phrased, the test directions, the categories for classifying the responses, the scoring criteria, and the validity criteria are all culture bound" (p. 579).

As noted previously, there are no tests that have norms suitable for individuals who are either bilingual or bicultural (or, frequently, both). Level of acculturation and English language proficiency have yet to be stratified along with other variables in the construction of norm samples. This is problematic because in the United States, whenever an individual who does not speak English or is limited in English proficiency enters the public school system, he or she will invariably be required to learn English, or in a sense become a circumstantial bilingual. Even when such individuals become "English dominant" or "fully English proficiency," they do not cease to be bilingual. Their dual-language backgrounds are not erased simply because they are now able to communicate at a conversational level and with no trace of accent. Even nonverbal tests, which reduce the inherent problems with communication and linguistic loading, do not address the issues of acculturation and the norms remain based primarily if not exclusively on monolingual English speakers. Native language tests also have the same problem. That is, although they tend to reduce bias that may be due to language comprehension or proficiency issues, they nevertheless are normed on monocultural, monolingual speakers of the native language—a group that is no more an appropriate standard for comparison than monocultural, monolingual English speakers when the examinee happens to be bilingual and bicultural.

The only practical recommendation that may be offered with respect to nondiscriminatory assessment is to first recognize that use of any test with diverse individuals will result in some form of bias but that the degree of bias can vary greatly. Initially, psychologists seemed intent on finding ways in which an actual IQ could be "adjusted" or modified so as to offer a better estimate of an individual's true functioning, often by taking into account acculturation issues as well as norm sample issues in estimating the eventual IQ. However, such practices, as well as others that have utilized nonstandardized administrations, altered norms, and the like, have been met with harsh criticism and have failed to provide a viable framework for nondiscriminatory assessment. A more promising practical approach to dealing with the potential bias in the use of tests seems to rest in simply recognizing that all tests carry differing levels of expectations regarding language ability, embedded cultural content,

or both. By selecting tests that have less cultural loading and lower linguistic demands, more confidence can be placed in the fact that what is being measured is more a function of actual ability than lack of acculturation or English proficiency. This is not to say that the entire process of assessment has been rendered equitable but, rather, that the discriminatory aspects and potential in the use of tests have been considered and accommodated to a certain degree.

An example of a systematic framework for engaging in this type of selection and interpretive process is available via use of the CHC Culture–Language Matrix. Initially, the matrix provides a visual representation of cognitive ability tests as classified according to the degree of cultural loading and linguistic demand that allows practitioners to select tests that are lower along these dimensions and thus more likely to be equitable or fair. The matrix also provides a method for evaluating actual vs expected patterns of performance that draws upon research in order to guide interpretations regarding the validity of the results relative to the primary or contributory effects of cultural or linguistic differences.

By allowing practitioners to select tests that may provide fairer estimates of true ability as well as a practical way to gauge the extent to which cultural or linguistic factors may have played a role in influencing test performance, potential bias and discrimination in test use can be effectively reduced, albeit never eliminated entirely. Nonetheless, when it can be ascertained that cultural or linguistic factors did not play a primary role in determining the obtained pattern of test results, a significant portion of the potential inequity (related to validity) has been resolved and practitioners can exercise more confidence in the validity of their results. Ultimately, nondiscriminatory assessment is not dependent on so-called culture-free or nonverbal tests but rather on a process that recognizes the nature and source of possible invalidity and that seeks to reduce potential bias to the maximum extent possible.

5. NONDISCRIMINATORY ASSESSMENT

Several recommendations have been provided in the literature for engaging in various forms of nondiscriminatory assessment. Although many discrete methods for reducing bias in assessment have been described throughout the years, much less attention has been paid to the development of a broad, comprehensive framework for nondiscriminatory assessment. Within the field of school psychology, Ortiz provided a framework that detailed a 10-step process for nondiscriminatory assessment. This process included guidelines and recommendations on how to assess and evaluate the learning ecology; assess and evaluate language proficiency; assess and evaluate opportunity for learning; assess and evaluate educationally relevant cultural and linguistic factors; evaluate, revise, and retest hypotheses; determine the need for and language(s) of assessment; reduce bias in traditional testing practices; utilize authentic and alternative assessment procedures; evaluate and interpret all data within the context of the learning ecology; and link assessment to intervention.

The framework delineated by Ortiz integrates salient and promising procedures and recommendations for nondiscriminatory assessment culled from research and practice. It is intended to be both a linear guide to assessment and a recursive model that allows a return to previous steps as may be necessary when newly collected data suggest alternative hypotheses. The framework can accommodate assessment activities at the individual and team levels and is particularly suited toward collaborative assessment in which members of an assessment team (including parents) can work together, sharing data and information and drawing conclusions or making decisions jointly. The need to consider multiple sources of information and multiple points of view in any assessment underlies the principles of nondiscriminatory assessment and serves to enhance the validity and ultimately the fairness and equity of the evaluation.

6. SUMMARY

Nondiscriminatory assessment should not be viewed as a single procedure, method, or test. Rather, fair and equitable assessment results from the integration of a wide range of approaches that together form the basis for drawing valid conclusions regarding the meaning of the collected data. Decisions regarding functioning and performance will likely prove to be less discriminatory in cases in which the sources of bias have been recognized a priori and systematic steps were taken to control their influence in both the collection and the interpretation of data. Nondiscriminatory assessment should not be seen merely as a search for nonbiased

tests—an impossibility—or a form of assessment that is applied only in cases of cultural or linguistic diversity. The provisions in IDEA and other guidelines for nondiscriminatory assessment do not assert that it is intended for select groups only but rather that it should be conducted in a way that ensures that all individuals are evaluated in the least discriminatory manner possible. As such, nondiscriminatory assessment is an approach to assessment that is equally applicable and appropriate to any individual and there is no reason it cannot be applied as such.

See Also the Following Articles

Acculturation ■ Cross-Cultural Psychology, Overview ■ Educational and Child Assessment ■ Intelligence and Culture ■ Intelligence Assessment ■ Personality Assessment ■ Psychological Assessment, Standards and Guidelines for ■ Psychometric Tests ■ Training, Cross-Cultural ■ Translation

Further Reading

Denbo, S., Beaulieu, L. M., & Coleman, J. L. (2002). *Improving schools for African American students: A reader for educational leaders*. Springfield, IL: Thomas.

Flanagan, D. P., & Ortiz, S. O. (2001). *Essentials of cross-battery assessment*. New York: Wiley.

Kavale, S. R., & Forness, K. A. (2000). *The efficacy of special education*. Washington, DC: American Association on Mental Retardation.

Ortiz, S. O. (2002). Best practices in nondiscriminatory assessment. In A. Thomas, & J. Grimes (Eds.), *Best practices in school psychology IV*. Washington, DC: National Association of School Psychologists.

Patton, J. M. (1998). The disproportionate representation of African Americans in special education: Looking behind the curtain for understanding and solutions. *Journal of Special Education*, *32*(1).

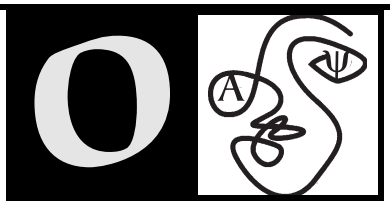
Sandoval, J., Frisby, C. L., Geisinger, K. F., Scheuneman, J. D., & Grenier, J. R. (1998). *Test interpretation and diversity: Achieving equity in assessment*. Washington, DC: American Psychological Association.

Sattler, J. (1992). *Assessment of children* (Rev. 3rd ed.). San Diego: Sattler.

Scarr, S. (1978). From evolution to Larry P., or what shall we do about IQ tests? *Intelligence*, *2*, 325–342.

Skiba, R. J. (2001). When is disproportionality discrimination? The overrepresentation of black students in school suspension. In W. Ayers, B. Dohrn, & R. Ayers (Eds.), *Zero tolerance: Resisting the drive for punishment in our schools* (pp. 176–187). New York: New Press.

Valdés, G., & Figueroa, R. A. (1994). *Bilingualism and testing: A special case of bias*. Norwood, NJ: Ablex.



Obesity

Shannon M. Kehle, Jessica S. Benas, and Michael A. Friedman

Rutgers, The State University of New Jersey, Piscataway, New Jersey, USA

1. Basic Knowledge of Obesity
2. Problems Associated with Obesity
3. Treatments
4. Future Directions
   Further Reading

GLOSSARY

binge eating disorder (BED) Recurrent episodes of eating large quantities of food while feeling a loss of control without the use of compensatory purging behaviors.

body image A person's concept of his or her physical image.

body mass index (BMI) A common measurement to assess obesity; BMI is the relation between a person's height and weight, expressed by weight in kilograms per height in meters squared.

obesity An accumulation of excess adipose tissue in relation to lean body mass.

reasonable weight loss A 5 to 10% loss of current weight; losses of this magnitude are achievable for most people and will lower their risk of developing obesity-related diseases.

stigma Being viewed as different from normative expectations based on cultural and situational factors; obese individuals who are stigmatized are viewed as possessing certain undesirable characteristics that are not necessarily accurate of every obese person.

toxic environment An environment in which high-fat, high-calorie foods are advertised heavily, offered at low costs, and readily available.

weight cycling Recurring instances of weight loss and regain, also known as yo-yo dieting.

Obesity is chronic condition that has recently reached epidemic proportions worldwide. Obesity is associated with serious medical and social consequences. Although the psychological consequences of obesity are less consistent, body image disturbances and binge eating disorder are common. While there are efficacious medical and behavioral treatments, maintenance of weight loss continues to be a problem and is a direction for future research.

1. BASIC KNOWLEDGE OF OBESITY

1.1. Definition

Obesity is defined as possessing a body mass index (BMI) of at least 30 kilograms per square meter (kg/m^2), and morbid/extreme obesity is defined as possessing a BMI of at least 40 kg/m^2. Besides using BMI, there are various ways in which to measure total body fat and classify obesity such as imaging techniques to assess localization of body fat, underwater weighing, skinfold thickness, and waist/hip circumference ratio. Obesity is often a chronic condition in which individuals who attempt to diet subsequently regain their lost weight, leading to "weight cycling."

1.2. Prevalence

Obesity is a health risk that has reached epidemic proportions worldwide, with more than 300 million adults

©2004 Elsevier Inc.
All rights reserved.

considered to be obese. Although prevalence estimates have differed due to various classification methods, recent statistics cite that nearly one-third of adult Americans are obese, a twofold increase over the past decade. Childhood obesity has also increased, with estimates suggesting that 15% of children and adolescents were obese in 2000, a 4% increase over the previous decade. In addition, there has been a marked increase in rates of morbid obesity; the prevalence rose threefold between 1990 and 2000.

1.3. Causes

Overall, obesity is the result of an individual consuming more calories than he or she expends, with an additional 3500 calories eaten leading to 1 pound gained. A complex interaction of environmental, genetic, and behavioral factors influences both consumption and expenditure. Factors related to consumption include food choices as well as food availability. For example, a diet high in fat will provide more energy than will carbohydrates or proteins and will produce fewer satiety signals, thereby disabling the individual's ability to suppress hunger cues. Currently, we live in a "toxic environment" such that the availability of unhealthy foods increases the likelihood that poor decisions are made. For example, there is an increasing number of commercials for calorie-dense foods, such as candy and high-sugar cereal geared toward children, and an abundance of easily accessible fast-food restaurants offering fatty foods at low costs.

Caloric expenditure is influenced by several factors, including resting metabolic rate and level of activity. Although a range of genetic factors most likely influences metabolism, the U.S. society is moving toward an increasingly sedentary lifestyle given that many Americans are engaging in less planned and lifestyle exercise. Fewer than one-third of Americans get the minimum recommended amount of exercise of light to moderate activity five times per week. Thus, as a society, Americans are consuming more and expending less, causing a rise in rates of obesity.

2. PROBLEMS ASSOCIATED WITH OBESITY

2.1. Medical Consequences

Obesity is strongly associated with diabetes mellitus; close to 50% of all individuals with type 2 diabetes have a BMI of at least 30 kg/m^2. This relation results from insulin's decreased ability to act at a cellular level when there is an excess of fatty tissue in the body. Obesity is also a risk factor for the development of coronary heart disease (CHD). Obesity contributes to the development of CHD through increasing rates of risk factors such as hypertension and increased cholesterol. Recently, an association between obesity and cancer has also been found. Obesity is associated with increased risk for certain cancers, including colon, endometrial, and breast cancers, and is also related to increased mortality rates among cancer patients. The causal mechanisms linking obesity to cancer are currently unknown.

Obesity is also associated with other complications such as disruption of sleep, difficulties with daily activities (e.g., climbing stairs, walking, bending, kneeling), increased levels of lower back pain, shortness of breath, and fatigue. As a result, impairments in daily business, work productivity, and role functioning are significantly more common. Overall, obesity accounts for approximately 300,000 deaths nationwide each year, and treatment costs were as high as $117 billion in 2000.

2.2. Social Stigma

Obese individuals are not only condemned for their appearance but also blamed for their condition. In a society that espouses a thin and lean body ideal, obese individuals are viewed as less attractive. Furthermore, body weight is viewed as controllable, leading obese individuals to be perceived as lacking the self-discipline necessary to maintain thinness. As a result, obese individuals are often thought to be lazy, gluttonous, and stupid. These assumptions have led to discrimination against the obese in areas such as employment, education, housing, and medical care.

2.3. Psychological and Behavioral Consequences

Efforts aimed at clarifying psychological consequences of obesity have provided inconsistent results. The obese population appears to be heterogeneous with respect to psychological functioning, with a subset of individuals experiencing psychological consequences of their weight and others exhibiting healthy mental functioning. To gain a more comprehensive understanding of the psychological correlates of obesity, researchers have suggested an approach focused on identifying risk factors that may explain the occurrence of negative

psychological consequences of obesity. Potential risk factors for psychological distress include being female, being adolescent, and being severely overweight. In addition, a subgroup of obese individuals who suffer more psychological distress are binge eaters. Binging refers to eating a large quantity of food (an larger amount larger than an individual would normally eat) in a discrete amount of time and feeling a loss of control. When individuals engage in binge eating more than twice a week and do not engage in compensatory purging behaviors (e.g., vomiting), they may be diagnosed with binge eating disorder (BED). Obese binge eaters tend to display higher levels of psychological distress and social dysfunction than do nonbinging obese individuals.

One area where obesity carries a more consistent risk for suffering is body image dissatisfaction. Body image concerns have been found to increase steadily with BMI, and body image disorders are thought to be present in approximately 10% of the obese population. Body image disorders often result in emotional distress due to preoccupation with appearance, negative self-worth, and social isolation.

3. TREATMENTS

3.1. Medical Approaches

There are two medical approaches currently available for the treatment of obesity: surgery and pharmacological treatments. Because of potentially life-threatening complications, surgery is recommended only for the morbidly obese (those with a BMI of at least 40 kg/m^2) or for those with a BMI of at least 35 kg/m^2 and a comorbid physical illness. Bariatric surgery produces its effects through either limiting the stomach's capacity or interfering with the digestion of food. Gastric bypass is widely accepted as the surgery of choice due to its low mortality rate and the ability to produce losses of up to 70% of excess body fat. Surgical treatments have also been shown to effectively improve type 2 diabetes, cardiac conditions, and quality of life among obese individuals. Although the initial effects of the surgery are drastic, weight regain is likely without significant behavior modifications; the average regain is 15% of initial excess weight.

Pharmacological treatment may be appropriate for individuals with a BMI of 30 to 40 kg/m^2 and for the morbidly obese who do not opt for surgical treatment. Although weight loss drugs have proven to be unsafe in the past, significant advances have been made in this area. There are currently two weight loss drugs that have been approved by the Food and Drug Administration: sibutramine and orlistat. Treatment with sibutramine increases satiety levels and produces average losses of approximately 5% of total body weight with a relatively low side effect profile. Orlistat blocks absorption of approximately 30% of dietary fat consumed daily and has weight loss effects approximately equal to those seen with sibutramine. Both drugs have been shown to be more efficacious when combined with dieting, and losses of more than 10% of original body weight are common. The long-term effects of pharmacological treatments are less promising. There is evidence that continued long-term use may lead to decreased efficacy, and discontinuation of the drugs is associated with weight regain.

3.2. Very Low-Calorie Diets

Very low-calorie diets (VLCDs) are defined as programs that allow fewer than 800 calories per day. The goal of such programs is to produce extremely rapid weight loss while maintaining muscle mass. This is generally accomplished through consuming small amounts of prepackaged foods and powdered supplements. VLCDs are appropriate only for those with a BMI of at least 30 kg/m^2 due to the muscle depletion that would occur in lower weight individuals. For the obese, VLCDs have been found to be effective in short-term weight loss and safe when used under medical supervision. Weight losses of approximately 20 to 25% of initial body weight typically occur within 16 weeks of treatment. However, individuals who use VLCDs find it difficult to maintain their losses over a 1-year period. When considering both the risk of malnutrition and the large level of weight regain following VLCD treatment, there is substantial doubt as to whether VLCDs are preferable to more moderate diets.

3.3. Behavioral Approaches

Behavioral weight loss (BWL) approaches, also known as cognitive–behavioral treatments, are currently the psychological treatment of choice for obesity. BWL treatment for obesity is generally conducted in a group format and occurs once per week for 4 to 6 months. The primary goals of BWL treatments are to increase physical activity levels and decrease caloric restriction so as to create a negative energy balance. Patients are encouraged to set reasonable health goals and follow moderate dieting plans that provide sufficient nutrition and allow for flexibility. Some of the core components of

BWL for obesity include (a) self-monitoring, (b) identification of factors that influence eating, (c) stimulus control, and (d) cognitive restructuring. Self-monitoring is an essential component of behavior change. Patients record factors such as their food intake, emotions while eating, and daily exercise. These records allow patients to identify and work toward changing behavior patterns. Once eating cues are discovered, stimulus control is used to limit exposure to such cues. Finally, cognitive restructuring is used to reshape attitudes about dieting, weight, and body image.

The effects of BWL have been widely studied. BWL has been shown to produce average weight losses of 7 to 10% of total body weight and to significantly improve both the medical and psychological consequences of obesity. However, as with other treatments, maintaining initial losses has proven to be difficult for patients. Approximately one-third of the losses are typically regained within 1 year, and slow weight gain tends to continue until baseline levels are met. It appears as though patients are unable to continue with the behavioral changes over long periods of time. One possible reason is that the degree of weight loss that patients have experienced is not rewarding enough to continue with the challenging lifestyle modifications.

3.4. Binge Eating Disorder

During the short time that BED has been recognized as a distinct eating disorder, several effective treatments have emerged. Two specific psychological treatments for BED have been tested: cognitive–behavioral therapy (CBT) and interpersonal therapy (IPT). The focus of CBT for BED is similar to the focus of CBT (BWL) for the nonbinging obese; however, an additional emphasis is placed on moderation of food intake so that the amount consumed is neither restrictive nor excessive. Initial studies of CBT for BED have shown that treatment cuts binge frequency in half and that approximately 60% of patients abstain from binging for 1 year. Self-help versions of CBT also show promising results, with abstinence rates approximately equal to those of group CBT treatments.

IPT for BED is a directive treatment that focuses on managing interpersonal difficulties that are thought to result in emotional states that trigger binge episodes and maintain problem behaviors. Specific eating behaviors are rarely discussed within treatment, and weight loss strategies are not included; eating and weight are addressed only as they relate to interpersonal issues. Controlled studies of the efficacy of IPT have revealed short-term and follow-up results approximately equal to those of CBT. Finally, BWL strategies as used in the treatment of obesity have been shown to be as efficacious as CBT and IPT in reducing binge frequency. In addition, BWL programs show an added benefit of weight loss that is not seen in CBT or IPT. For this reason, BWL programs are considered by many to be the preferred treatment for BED.

3.5. Strategies for Weight Loss Maintenance

Effective weight loss treatments are available; however, the maintenance of losses continues to be a challenge. Three primary strategies aimed at increasing maintenance rates have been proposed recently. The first involves combining medical approaches with behavioral lifestyle modifications. It has been shown that combining surgery with psychological approaches aimed at healthy eating and exercise may be effective in maintaining initial losses. Also, the addition of lifestyle modification has been shown to more than double the effects of drugs alone and to prevent weight regain during the maintenance phase of treatment. Therefore, to provide optimal treatment, behavioral treatments need to be considered a standard adjunct to pharmacological interventions.

The second strategy is extending the length of existing CBT treatments. This has been shown to result in continued adherence to behavioral changes and, in turn, to improved rates of maintenance. The focus of the extended contact would be teaching problem-solving techniques that could be used in challenging situations and stressing the importance of maintaining behavioral changes. Although such strategies have been proven to be efficacious, attrition rates following initial weight loss are high; therefore, interventions aimed at increasing retention rates are necessary.

The third proposed strategy is to modify patients' expectations regarding the amount of weight loss that can be expected. For example, current CBT weight loss programs are typically able to produce losses of approximately 10% of initial body weight. Although a loss of 10% substantially improves health-related quality of life and medical complications related to obesity, patients typically expect greater weight losses and are disappointed by modest results. It has been proposed that the dissatisfaction experienced with modest losses, coupled with the effort required to maintain such losses, leads to an abandonment of behavioral changes and subsequent weight regain. Therefore, experts are

now suggesting that patients aim for modest losses rather than "ideal" losses. A new form of CBT for obesity that includes the modification of beliefs regarding the desired amount of weight loss and places an emphasis on acceptance of modest losses is currently being tested. Although realistic expectations regarding weight loss may increase maintenance rates, convincing obese individuals to engage in treatment for minimal losses will undoubtedly prove to be difficult.

4. FUTURE DIRECTIONS

Given the growing prevalence of obesity worldwide, the demand for effective treatments continues to increase. In response, researchers are continually working to improve existing treatments and develop new approaches. Medical breakthroughs will undoubtedly continue, and new psychological treatments are currently being developed and evaluated. However, obesity likely will remain a difficult disorder to treat. Brownell recently argued that a preventive approach, aimed at combating the toxic environment that supports obesity, is necessary to reverse the obesity epidemic. Brownell suggested that a public health approach with government initiatives, such as taxing fattening foods, removing unhealthy foods from public schools, and providing federal funding for nutrition programs, will be necessary to eliminate the environment that supports overeating and inactivity. Until such steps are taken, the ability of weight loss treatments to produce enduring change may remain limited. Public health policy might prove to be necessary in preventing obesity and maintaining long-term weight loss.

See Also the Following Articles

Health and Culture

Further Reading

Agras, W. S., Telch, C. F., Arnow, B., Eldredge, K., & Marnell, M. (1997). One-year follow-up of cognitive behavioral therapy for obese individuals with binge eating disorder. *Journal of Consulting and Clinical Psychology*, *65*, 343–347.

Allison, D. R., Fontaine, K. R., Manson, J. E., Stevens, J., & Van Itallie, T. B. (1999). Annual deaths attributable to obesity in the United States. *Journal of the American Medical Association*, *27*, 1530–1538.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Brownell, K. D. (1991). *The LEARN program for weight control*. Dallas, TX: American Health Publishing.

Brownell, K. D. (2004). *Food fight*. New York: McGraw–Hill.

Carter, J. C., & Fairburn, C. G. (1998). Cognitive–behavioral self-help for binge eating disorder: A controlled effectiveness study. *Journal of Consulting and Clinical Psychology*, *66*, 616–623.

Centers for Disease Control and Prevention. (2003). *Defining overweight and obesity*. www.cdc.gov.

Fairburn, C. G., & Brownell, K. D. (Eds.). (2002). *Eating disorders and obesity* (2nd ed.). New York: Guilford.

Fairburn, C. G., & Wilson, G. T. (Eds.). (1993). *Binge eating*. New York: Guilford.

Foster, G. D., Wadden, T. A., Vogt, R. A., & Brewer, G. (1997). What is a reasonable weight loss? Patients' expectations and evaluations of obesity treatment outcomes. *Journal of Consulting and Clinical Psychology*, *65*, 79–85.

Friedman, M. A., & Brownell, K. D. (1995). Psychological correlates of obesity: Moving to the next research generation. *Psychological Bulletin*, *117*, 3–20.

Gortmaker, S. L., Must, A., Perrin, J. M., Sobol, A. M., & Dietz, W. H. (1993). Social and economic consequences of overweight in adolescence and young adulthood. *New England Journal of Medicine*, *329*, 1008–1012.

Jeffery, R. W., Drewnowski, A., Epstein, L. H., Stunkard, A. J., Wilson, G. T., Wing, R. R., & Hill, D. R. (2000). Long-term maintenance of weight loss: Current status. *Health Psychology*, *19*, 5–16.

National Institute of Health. (2003). *Statistics related to overweight and obesity*. www.niddk.nih.gov.

National Task Force on the Prevention and Treatment of Obesity. (1993). Very low-calorie diets. *Journal of the American Medical Association*, *270*, 967–974.

Perri, M. G., Nezu, A. M., Patti, E. T., & McCann, K. L. (1989). Effect of length of treatment on weight loss. *Journal of Consulting and Clinical Psychology*, *57*, 450–452.

Pi-Sunyer, F. X. (1993). Medical hazards of obesity. *Annals of Internal Medicine*, *119*, 655–660.

Prochaska, J. O., & DiClemente, C. C. (1983). Stages and processes of self-change of smoking: Toward an integrative model of change. *Journal of Consulting and Clinical Psychology*, *51*, 390–395.

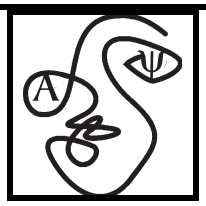
Puhl, R., & Brownell, K. D. (2003). Ways of coping with obesity stigma: Review and conceptual analysis. *Eating Behaviors*, *4*, 53–78.

Rossner, S. (2002). Obesity: The disease of the twenty-first century. *International Journal of Obesity*, *26*, S2–S4.

Wadden, T. A., Berkowitz, R. J., Sarwer, D. B., Prus-Wisniewski, R., & Steinberg, C. (2001). Benefits of lifestyle modification in the pharmacologic treatment of obesity. *Archives of Internal Medicine*, *161*, 218–227.

Wadden, T. A., Brownell, K. D., & Foster, G. D. (2002). Obesity: Responding to the global epidemic. *Journal of Consulting and Clinical Psychology*, *70*, 510–525.

World Health Organization. (2003). *Nutrition*. www.who.int.



Occupational Choice

Jo-Ida C. Hansen and Sarah K. Stoever
University of Minnesota, Minneapolis, Minnesota, USA

1. Career Theories
2. Career Counseling for Occupational Choice
3. Conclusion
   Further Reading

GLOSSARY

maxi-cycle A series of life stages characterized as a sequence of growth, exploration, establishment, maintenance, and disengagement.
satisfaction A perception of satisfaction from the perspective of the employee.
satisfactoriness A perception of satisfaction from the perspective of the employer.
Self-Directed Search A self-administered inventory that helps people to explore what occupations to follow and to estimate related abilities.
Strong Interest Inventory A self-report inventory that inquires about interests in a wide range of occupations, occupational activities, hobbies, leisure activities, school subjects, and types of people.

What makes people who are big lottery winners keep their jobs even if they are in low-paying positions? Of 16 women in a small town (population 736) in Minnesota who won $95.5 million (approximately $2 million each after taxes), 15 claimed that the money was not reason enough to quit their morning jobs as school cafeteria workers or their second jobs driving school buses, cleaning houses, or milking cows. And what about the 16th woman? At 70 years of age, she already was considering retirement. Their story is not unusual. In 2004, Arvey and colleagues found that few lottery winners quit working completely, although some do take sabbaticals or change positions. These women certainly did not stay in their jobs due to the high wages they earned; their pay ranged from $7.75 to $10.68 per hour. Rather, research shows that a variety of internal and external factors keep people on the job—some grudgingly, of course, but others with passion and zeal for what they do. Obviously, people who are not big lottery winners do work for financial reasons—to meet basic needs or to be able to afford nonwork activities that cost money. But beyond income, work gives people a feeling of belonging and provides structure. It gives people a sense of satisfaction and productivity while satisfying their interests, promoting their self-esteem, and providing them with status and recognition. Work also provides a personal identity and a mechanism for describing oneself to others. Often, what people do is who they are. Most adults spend one-third to one-half of their waking hours at work. It follows, then, that work can have a tremendous impact on psychological and physiological well-being. It is little wonder that the process of making occupational choices can create anxiety as one searches for the perfect match.

©2004 Elsevier Inc.
All rights reserved.

1. CAREER THEORIES

Since the early 1900s, scholars have studied occupational choice and career decision making. They focused first on improving personal satisfaction, and later on increasing performance and lowering turnover, through informed occupational choices. Several theories have emerged that identify the variables important to the process of occupational choice and work adjustment. The theory of work adjustment focuses on the congruence between the abilities and needs of the individual and the requirements and rewards of the job, Super's life span theory focuses on the developmental process of occupational choice, and Holland's theory of vocational personality type focuses on prediction of the types of work that people will choose. These three theories complement one another and offer a framework for understanding occupational choice and adjustment. Career theories also provide a mechanism for conceptualizing career concerns and client goals. For example, the theory of work adjustment can guide conceptualizing clients with adjustment problems in the workplace, Super's theory can lend understanding to the influence of developmental stages in occupational choice, and Holland's theory can provide a model for describing the world of work.

1.1. Theory of Work Adjustment

The theory of work adjustment, stemming from the Work Adjustment Project at the University of Minnesota, first appeared during the mid-1960s and has subsequently been revised and extended a number of times. The theory describes the dynamic interaction between persons and their work environments that influences work adjustment and occupational choice.

According to Dawis and Lofquist, the architects of the theory of work adjustment, individuals "inherently seek to achieve and maintain correspondence with their environment," where correspondence is defined as a well-balanced relationship between the characteristics of the individual and the requirements of the environment. Indeed, much of the emphasis of the theory concerns the matching of the abilities and values of an employee with the abilities required by a job and the reinforcers provided by the job. Consequently, the theory provides a model for career counseling that is designed to help individuals achieve (a) a clear understanding of their own abilities, values, personalities, and interests; (b) a knowledge of the requirements and conditions of success in a work environment; and (c) a true understanding of the interaction between these two groups.

Dawis and Lofquist maintained that abilities are general dimensions that underlie groupings of required skills, including general learning ability, verbal ability, numerical ability, spatial ability, form perception, clerical ability, eye–hand coordination, finger dexterity, and manual dexterity. Similarly, values are seen as representing a grouping of needs and are defined as achievement, comfort, status, altruism, safety, and autonomy. Two additional components to the prediction of work adjustment and occupational choice are satisfaction (i.e., a perception of satisfaction from the perspective of the employee) and satisfactoriness (i.e., a perception of satisfaction from the perspective of the employer). Put another way, satisfaction refers to the extent to which employees' needs are fulfilled by the work they do, whereas satisfactoriness concerns the appraisal by others, usually employers, of the extent to which the employees adequately complete the work that is assigned to them.

The theory of work adjustment is heavily based on the premise that individuals differ in their abilities, needs, values, and interests, among other variables. This principle views people as complex individuals who differ on a number of dimensions and suggests that counseling needs to be individualized to allow everyone to make a personalized occupational choice. Although the theory is broader than many career theories and has important implications for career adjustment and occupational choice, it focuses on only one point in time. Super's life span theory, on the other hand, focuses on the developmental process of occupational choice.

1.2. Super's Life Span Theory

Probably no one has written as extensively about career development or influenced the study of the topic as much as has Super. His earliest theoretical statements were influenced by researchers in various areas of psychology, sociology, and personality theory. In fact, Super described his theory of career development as a "segmental" theory or one that includes the work of many other theorists such as Freud, Maslow, and Rogers. From the works of these theorists, Super derived basic assumptions that allowed him to develop his own developmental process of occupational choice.

A hallmark of Super's theory is the view that occupational development is a process of making several decisions over time, culminating in occupational

choices that represent an implementation of the self-concept. Super viewed self-concept as a combination of biological characteristics, social roles people play, and evaluations of the reactions that others have toward the individual. For Super, self-concept served to organize roles that people play throughout their lifetimes.

An important idea throughout Super's career development theory is the concept of role. Super described six major roles—homemaker, worker, citizen, leisurite, student, and child—and stressed that these roles, as well as the importance of these roles, vary throughout one's life. For instance, he argued that during childhood, the roles of leisurite, student, and child are particularly important, while the roles of worker, citizen, and homemaker are minimal. During adolescence, citizen and worker may become more important roles, but work is not often directly related to one's eventual career. It is during adulthood that one has more occupational choices. Consequently, it is important that professionals understand the concept of life roles (in conjunction with developmental stages) when helping individuals to make occupational choices.

The notion of developmental stages is essential to Super's life span theory. Super proposed a series of stages or "maxi-cycles" over the life span, beginning with the growth stage during early childhood. During this stage, children become curious about life and begin to explore their environments. The information gained through this exploration will be an important component of the children's self-concept. The exploratory stage begins with adolescents' awareness that an occupation will be an aspect of life and ends with young adults choosing occupations that they believe are within reach and provide important opportunities. The establishment stage, as one would expect, relates to adults' early encounters within actual work experiences. During the maintenance stage, employees attempt to continue or improve their occupational situations. The final stage, disengagement, includes the preretirement period during which employees' emphasis on work is focused on retaining their positions rather than enhancing them. This period ends when older adults withdraw from the world of work.

Super proposed characteristic developmental tasks within each stage. Mastery of these tasks allows individuals to function effectively in their life roles within that stage and prepares them for the next task. Successful coping with the requirements of each stage is dependent on the individual's career maturity—a group of physical, psychological, and social characteristics that represent the individual's readiness and ability to deal with the developmental problems and challenges that are faced. The person whose maturity is equal to the problem probably resolves it with minimal difficulty or concern. However, when the maturity is not sufficient for the task, responses of procrastination, ineptness, and/or indecision are likely to occur.

As was mentioned previously, a fundamental aspect of Super's theory is that occupational choice is an implementation of the self-concept. The implication for career counselors is that, in addition to providing clients with information about who they are, counselors need to integrate objective information about the self (e.g., interests, values, abilities). Holland's theory of vocational personality type focuses on this objective information.

1.3. Holland's Theory of Vocational Personality Types

Holland's theory continues to enjoy attention as it celebrates more than four decades of popularity. Holland's person–environment typology and theory of career choice clearly is the most widely studied career theory in history. In addition, the concepts that derive from this theory are integral to the vocabulary, tools, and processes of career counseling.

Holland's theory of occupational choice is based on several assumptions. One assumption is that persons and environments can be categorized according to six types: Realistic, Investigative, Artistic, Social, Enterprising, and Conventional.

1. Realistic individuals are likely to enjoy using tools or machines in their work and often approach problems in a practical or problem-solving manner. Likewise, a Realistic environment is one in which the work setting includes tools, machines, and/or animals that the individual manipulates and where the ability to work with "things" is more important than the ability to interact with others.

2. Investigative persons are likely to enjoy challenges that require the use of intellect and are apt to enjoy learning and feelings of confidence about their ability to solve mathematical and scientific problems. Similarly, an Investigative environment is one in which people search for solutions to problems through mathematical and scientific interests and competencies.

3. Artistic individuals like the opportunity to express themselves in a free and unsystematic way by creating music, art, and/or writing. Correspondingly, an Artistic environment is one that is free and open and is encouraging of creativity and personal expression.

4. Social persons are interested in helping people through teaching, helping with personal or vocational problems, and/or providing personal services. Similarly, a Social environment is one that encourages people to be flexible and understanding of each other and where people can work with others through helping with personal or career problems, teaching others, and being socially responsible.

5. Enterprising individuals enjoy being with others and like to use verbal skills to sell, persuade, and/or lead. Correspondingly, an Enterprising environment is one in which people manage and persuade others to attain organizational or personal gains.

6. Conventional persons are typically dependable and have the ability to follow rules and orders in unambiguous situations. Likewise, a Conventional environment is one in which organization, planning, and ability to follow directions (often in an office environment) are of utmost importance.

Another assumption, coined Holland's calculus hypothesis, describes the relationship among the six person–environment types. The calculus hypothesis states that the six types can be arranged around a hexagon, with the types most similar to each other (e.g., Social and Enterprising) falling next to each other and those most dissimilar (e.g., Social and Realistic) falling directly across the hexagon from one another.

Holland's theory has been applied extensively in career counseling, most notably using the six-category typology to categorize individuals in the interpretation of interest inventories such as the Self-Directed Search (SDS) and the Strong Interest Inventory (SII). In addition, Holland's typology has been used to classify occupational information and college majors in an attempt to facilitate individuals' success in making occupational decisions.

1.4. Summary of Theories

The theory of work adjustment, Super's life span theory, and Holland's theory of vocational personality types all have received a great deal of attention and empirical support over the years. Research on the theory of work adjustment has generated many promising findings; however, research specifically using constructs from this theory has proceeded at a slow pace. Super's life span theory is intuitively appealing and offers assistance in understanding the richness of an individual's career and life. Nevertheless, it remains to be seen whether it will maintain its position of preeminence in the future. Finally, Holland's theory of vocational personality types stands as the most influential of the existing theories in both research and practice, and research supporting Holland's theory is very robust.

2. CAREER COUNSELING FOR OCCUPATIONAL CHOICE

Career theories suggest that people make their occupational decisions in an orderly and systematic fashion. Yet research shows that for most people, the path often is serpentine. On the one hand, barriers to occupational entry, such as stereotyping, prejudice, and family, economic, and educational constraints, thwart people's opportunity to follow their interests. On the other hand, roughly 60% of adults report that serendipity plays a role in helping them to make good career decisions. Career counseling is a process, facilitated by a counselor, designed to help people find their way through the decision-making process. The ultimate goal is to help people obtain satisfaction and effectiveness in their jobs, occupations, and careers.

Career counseling can be structured in several different delivery modes. In some cases, the intervention is one-on-one between the client and the counselor. In other instances, the intervention may be conducted for groups and may resemble an educational experience with assignments and "in-class" discussions and information sharing. Career counseling also can occur in a variety of settings. For example, it is not unusual for college students to seek career counseling to help them decide on majors or to help them focus their job searches at the time of graduation. Many high schools also offer students the opportunity to explore future careers either with their school counselors or in life studies courses. Some employers offer career exploration opportunities within the work setting or arrange referrals for clients to work with career counselors. In addition, community organizations (e.g., the YWCA/YMCA) and adult education centers offer career counseling. Many people also seek out career counselors (now sometimes called career coaches) who are in private practice or work for consulting firms.

The desire to learn more about oneself and to ponder career questions is not the privilege of only the young. Many adults turn to career counseling for a variety of reasons. Some people are genuinely unhappy in their jobs and seek career counseling with the goal of finding better matches for their interests,

personalities, and values. Some people are bored and believe that their abilities are underused or that they have reached career plateaus. Others, in spite of their achievements, may question whether they are making a difference. Still others believe that they drifted into their jobs and never made deliberate choices. Not all people will discover, even with the help of career counseling, that they have a passion for work. For some people, work is necessary simply to provide for their basic needs, and other aspects of their lives will give them more pleasure and satisfaction than will their work. However, career counseling can help even those not driven to career achievement to maximize the match between their personal characteristics and those of the jobs they choose.

Regardless of the format, setting, life stage, or goals and purposes, most career counseling approaches involve psychological assessment. The ideal situation allows assessment of interests, abilities, personality, values, and needs. However, time and financial constraints often limit testing to the administration of a vocational interest inventory such as the SII, the Campbell Interest and Skill Survey (CISS), the SDS, the Career Assessment Inventory (CAI), or the Sixteen Personality Factor Questionnaire (16PF) (Table I).

The University of Minnesota's Vocational Assessment Clinic (VAC) offers one model for combining comprehensive assessment with career planning services during a time-limited intervention: intake, assessment, interpretation of inventories, and integration with plan of action. In addition to taking the SII to assess vocational interests, clients who use the VAC services take the Minnesota Ability Test Battery (MATB), the California Psychological Inventory (CPI), and the Minnesota Importance Questionnaire (MIQ). During the interpretive sessions, counselors use the test results to help structure career exploration, to stimulate new possibilities and directions, and generally to help clients develop an integrated whole of their self-knowledge. Interest inventories can provide clients with a wealth of information, and the number of possibilities to explore often can be overwhelming. Thus, additional information from the assessment of abilities, personality, values, and needs can be useful for focusing their exploration and for providing additional information to determine the extent to which their characteristics match those of various jobs and occupations. After exploring the assessment results, the counseling process moves to developing a plan of action, often in the form of outside assignments, for gathering more information to inform the career decisions.

TABLE I
Comparison of Interest Inventories

Vocational interest inventory	*Focus*	*Usage setting(s)*	*Number of items*	*Scales*
Strong Interest Inventory	Professional populations Separate sex norms	High school/College	317	6 General Themes 25 Basic Interest 211 Occupational 4 Personal Style
Self-Directed Search	Student populations Self-administered Self-scored	High school/College	228	Daydreams Activities Competencies Occupations Self-Estimates
Campbell Interest and Skill Survey	Professional populations Combined sex norms Addresses self-estimated skills	Business	320	7 Orientation 29 Basic Interest 60 Occupational 2 Special
Career Assessment Inventory	Nonprofessional populations Combined sex norms	High school/College	370	6 RIASEC Themes 25 Basic Interest 111 Occupational 4 Nonoccupational
Sixteen Personality Factor Questionnaire	Professional populations Addresses personality	Business	185	16 Primary Factors 5 Global Factors 3 Response Styles

3. CONCLUSION

Research has been done to determine the effectiveness of career counseling. The results across several studies indicate that an average client has an outcome that is better than the outcomes of 80% of untreated control groups. Research has shown that congruence between a person's interests, abilities, personality, and values and the work environment—a desirable outcome of career counseling—leads to success and performance on the job and in school, life and work satisfaction, persistence on the job (i.e., tenure), and positive self-image and self-esteem. Conversely, poor correspondence (i.e., person–environment incongruence) is related to burnout, increased levels of anxiety, and somatic complaints. The implication of these outcome results is that whether people view their work as a job, a career, or a calling, career counseling can be a powerful aid in making occupational choices that lead to well-being.

See Also the Following Articles

Career Counseling ■ Holland's Theory (Vocational Personality Types) ■ Indecision, Vocational ■ Job Analysis, Design, and Evaluation ■ Job Search ■ Vocational Interests ■ Vocational Psychology, Overview

Further Reading

Arvey, R. D., Harpaz, I., & Liao, H. (in press). Work centrality and post-award work behavior of lottery winners. *Journal of Psychology*.

Betsworth, D., & Hansen, J. C. (1996). The categorization of serendipitous career development events. *Journal of Career Assessment*, *4*, 91–98.

Brown, D., & Brooks, L. (Eds.). (1996). *Career choice and development* (3rd ed.). San Francisco: Jossey–Bass.

Campbell, D. P., Hyne, S. A., & Nilsen, D. L. (1992). *Manual for the Campbell Interest and Skill Survey*. Minneapolis, MN: National Computer Systems.

Cattell, R. B., Cattell, A. K., & Cattell, H. E. (1993). *Sixteen Personality Factor Questionnaire* (5th ed.). Champaign, IL: Institute for Personality and Ability Testing.

Dawis, R. V., & Weiss, D. (1994). *The Minnesota Ability Test Battery technical manual*. Minneapolis: University of Minnesota, Vocational Psychology Research.

Gough, H. (1995). *California Psychological Inventory manual*. Palo Alto, CA: Consulting Psychologists Press.

Hansen, J. C., & Dik, B. (2004). Measures of career interests. In M. Hersen, & J. Thomas (Eds.), *Handbook of psychological assessment*, Vol. 4: *Industrial/Organizational assessment*. New York: John Wiley.

Harmon, L., Hansen, J. C., Borgen, F., & Hammer, A. (1994). *Strong Interest Inventory applications and technical guide*. Stanford, CA: Stanford University Press.

Holland, J. L., Powell, A. B., & Fritzsche, B. A. (1994). *The Self-Directed Search professional user's guide*. Odessa, FL: Psychological Assessment Resources.

Judge, T. A., & Watanabe, S. (1993). Another look at job satisfaction–life satisfaction relationship. *Journal of Applied Psychology*, *78*, 939–948.

Leong, F. T. L. (Ed.). (1995). *Career development and vocational behavior of racial and ethnic minorities*. Hillsdale, NJ: Lawrence Erlbaum.

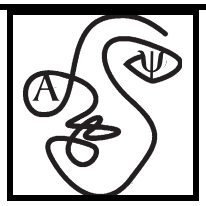
Meir, E. I., Melamed, S., & Dinur, C. (1995). The benefits of congruence. *Career Development Quarterly*, *43*, 257–266.

Parsons, F. (1909). *Choosing a vocation*. Boston: Houghton Mifflin.

Sharf, R. S. (2002). *Applying career development theory to counseling* (3rd ed.). Pacific Grove, CA: Brooks/Cole.

Strong, E. K., Jr. (1943). *Vocational interests of men and women*. Stanford, CA: Stanford University Press.

Ton, M., & Hansen, J. C. (2001). Using a person–environment fit framework to predict satisfaction and motivation in work and marital roles. *Journal of Career Assessment*, *9*, 315–331.



Occupational Psychology, Overview

Neil Anderson

University of Amsterdam, Amsterdam, The Netherlands

1. Introduction
2. Theoretical Perspectives
3. Key Topics in Occupational Psychology
   Further Reading

GLOSSARY

contracts Employment and psychological contracts form the basis of the work relationship between employers and employees. Employment contracts tend to be in written form, stipulating legally binding terms and conditions of work. Psychological contracts are sets of shared understandings, obligations, and expectations of the work relationship.

job stress and burnout The negative psychological effects of workplace demands relative to the personal resources of the individual to cope with environmental stimuli. Stress is the interaction between job environment stimuli and individual responses in situations where the individual does not possess sufficient resources to cope. Burnout is characterized by feelings of emotional exhaustion, depersonalization, a lack of personal accomplishment, and a manifest inability to cope with even basic, ongoing demands in the work environment.

occupational psychology The scientific sub-discipline concerned with individual, work team, and organizational performance, psychological well-being, and human integration in the workplace.

occupation and gender Topics concerned with occupational choice, socialization, job performance, career choice, and occupational choice as they relate to gender differences between males and females in the work environment.

personal initiative and innovation Topics concerned with idea generation, applied creativity, and innovation implementation in the workplace. Personal initiative involves work role changes that are self-initiated, that are proactive in orientation, and that demand persistence until change is implemented. Creativity is the generation of new ideas, whether eventually implemented or not. Innovation is the generation and implementation of new ideas, products, or processes at the individual, workgroup, or organizational levels of analysis.

personnel selection The attraction, recruitment, psychological assessment, and decision-making process involved in selecting newcomers for job vacancies in an organization in a valid, reliable, fair (i.e., legally defensible), and cost-effective manner.

work team A collective of two or more individuals that performs organizationally relevant tasks, shares one or more common goals, interacts, exhibits task interdependencies, manages boundaries, and is embedded in a broader organizational context.

Occupational psychology is a discipline related to industrial-organizational psychology that is concerned with individual, work team, and organizational performance, psychological well-being, and human integration in the workplace. Used especially in the United Kingdom, the term occupational psychology is used synonymously with that of industrial-organizational (I/O) psychology in the United States and Australia, and

©2004 Elsevier Inc.
All rights reserved.

with that of work and organizational (W/O) psychology in continental Europe. Due to its historic background in the United Kingdom, the subfield of occupational psychology developed somewhat specific foci, methods, and research concerns. Here, however, the term is used in its more general sense with acknowledgement to these historic precedents. Occupational psychology concerns work performance outcomes and health and well-being outcomes at three interrelated levels of analysis: (1) individual, (2) work group or team, and (3) organizational. Additionally, a few authors suggest a fourth level of analysis, that of interorganizational outcomes, although this has been more the concern of management scientists and organizational sociologists.

1. INTRODUCTION

1.1. Domain-Relevant Fields

The British Psychological Society (BPS) identifies eight major areas of knowledge and professional practice as constituting occupational psychology: (1) human–machine interaction, (2) design of environments and of work, (3) personnel selection and assessment, (4) performance appraisal and career development, (5) counseling and personal development, (6) training, (7) employee relations and motivation, and (8) organizational development. For individuals to become a chartered occupational psychologist (COP) under the BPS framework (essentially equivalent to licensing in other countries), they must have knowledge of all eight areas, supervised practice in five areas, and higher level independent practice over several years in at least two areas. Ongoing continuing professional development (CPD) workshops and seminars ensure that practitioners are kept updated on recent developments, but tensions between practice and research have risen under this framework. Researchers pursuing a Ph.D. and subsequent scientific career have found it difficult to gain the breadth of practical experience necessary to become COPs, as their work is usually specialized within one of the eight major areas. Recently, however, the Occupational Division of the BPS has issued tenders for project work to re-examine and possibly update this framework.

Arnold *et al*. stipulated 12 domains of occupational psychology, with some overlaps with the BPS professional society listing: (1) selection and assessment, (2) training, (3) performance appraisal, (4) organizational change and development, (5) ergonomics and equipment design, (6) career choice, development and counseling, (7) interpersonal skills, (8) equal opportunities, (9) occupational safety and health, (10) work design, (11) attitude surveys, and (12) well-being and work. Again, the overlaps with the domain knowledge frameworks used to train I/O psychologists in the United States and Australia, and with European syllabi for W/O psychologists stipulated by the European Association of Work and Organizational Psychology (EAWOP) and the European Federation of Psychologists' Associations (EFPA), are self-evident. Occupational psychology, nevertheless, has several areas of interest and active research that are not widely represented in other countries. These include occupational choice and career guidance, gender studies, health and safety in the workplace, including accidents and errors, and historically, unemployment research, especially the negative psychological effects upon individuals of being jobless.

1.2. Brief History

Within the context of the United Kingdom, occupational psychology has a history of some 50–60 years. Industrial psychology, as it was called then, began during World Wars I and II as a highly focused effort to support the armed services and government departments. The inter-war years also saw the creation of the National Institute for Industrial Psychology (NIIP) and the Medical Research Council, which subsequently provided funding to several research programs in the United Kingdom, most notably the applied psychology unit at Cambridge University and the social and applied psychology unit at Sheffield University under Peter Warr (later, under Toby Wall). These all developed from war-time work and research studies into such topics as soldier selection and "shell shock" among soldier casualties, and demobilization studies into how best to return soldiers and commissioned officers to civilian and commercial careers in peacetime. This period also witnessed formative visits to Britain by U.S. and European applied psychologists, including, most influentially, Viteles in 1922.

The inter-war years brought with them troubles for occupational and applied psychologists. The economic depression and the heavy reliance of the NIIP upon industrial funding caused Elton Mayo to decline the position of its directorship upon C.S. Myers' retirement in 1938. His rationale for this decision is fascinating, describing the reliance for research funding upon industrial sources as being "made at the request of troubled firms" and resulting in "planlessness" of such

an unprogrammatic approach to research. Exactly the same ills bedeviled occupational psychology and I/O psychology internationally in later periods of economic depression, when research funding was exclusively reliant upon business sources and commercially applied studies. Yet the NIIP was highly influential; the titles of some of its published reports closely resembled the domain areas noted above, which were to become the main foci of occupational psychology (e.g., *Industrial Accidents, Vocational Guidance and Selection, Methods of Work, Psychological Effects of Noise, Labour Turnover, and Hours of Work*). The NIIP and, later, the Tavistock Clinic formed the foundation of much commercial and applied psychology in Britain, but additionally the three armed services (Army, Navy, and Air Force) committed substantial resources, during World War II and after, toward research on occupational issues, including the development of soldier selection tests and exercises. Work by F.C. Bartlett at the Cambridge Psychological Laboratory (later named the Applied Psychology Unit) also contributed to these themes, again with a distinguishing feature being the synthesis between research on selection and psychological health outcomes. Post-war applications included the development of assessment centers for the British government (the so-called Civil Service Selection Board, or CSSB, which had close parallels to U.S. assessment center work at that time), pilot selection and error, naval officer selection and training, and pilot ergonomic studies.

The 30-year period between 1960 and 1990 witnessed substantial growth in both the numbers of registered occupational psychologists and the range of activities they were engaged in. The establishment of the Careers Service to guide school leavers in labor market entry, research work conducted at the Tavistock Institute (the so-called human relations movement), quality of working life (QWL) studies, and industrial training studies in the 1960s and 1970s culminated in the Socialist-Labour government of the time creating the Steering Group on Job Satisfaction, with representatives from trades unions and employers, and chaired personally by the Minister of State. The economic boom years of the 1980s led to a concomitant growth in occupational psychology, with COP numbers rising substantially. Consultancy firm growth was also a marked feature of this period, with firms such as SHL, Oxford Psychologists Press, and the Psychological Corporation expanding rapidly. International diversification by consultancies led several large U.S. firms to establish European subsidiaries, and also led SHL to set up offices in the United States. Growth spurred a step-shift in the numbers of accredited occupational psychology masters degree programs in the United Kingdom from a mere handful at the start of the 1980s to almost 30 at the end of the 1990s; in fact, the accreditation procedure itself was introduced by the BPS Training Committee during this period. Health and human resource management (HRM) issues continued to be high on the agenda for research and consultancy, stimulated still further at the end of the century by legal precedent that allowed employees to sue their organizations for claimed negative psychological effects caused by allegedly negligent HRM practices.

1.3. Defining Occupational Psychology

Modern-day use of the term occupational psychology has come to be synonymous in the United Kingdom with the terms I/O psychology in the United States and W/O psychology in continental European countries. Several handbooks and textbooks offer useful overview definitions, for example, "The application of psychological principles, theory, and research to the work setting" (Landy and Conte, 2004, p. 6), and "Work psychology is about people's behavior, thoughts and emotions related to their work. It can be used to improve our understanding and management of people" (Arnold *et al.*, 2004, p. 10).

1.4. Applications of Occupational Psychology

As psychology applied to the world of work, occupational psychology has generated a whole host of practical applications. This science–practice blend is both the essence of the discipline and cause for some periodic tensions. Porteous noted 13 main areas of application: (1) individual differences, (2) personality and work, (3) work design and measurement, (4) performance appraisal, (5) behavior and work performance, (6) humanistic values, (7) cross-cultural values, (8) counseling workers, (9) stress, (10) leadership, (11) human–computer interaction, (12) health and safety, and (13) new work. Occupational psychologists are employed in industry, government, and consultancies if functioning as practitioners, and by universities or research institutes if working as research scientists.

2. THEORETICAL PERSPECTIVES

As with other areas of applied psychology, occupational psychology has adopted a range of theoretical perspectives and approaches as it has developed over the decades. Distinct from other areas of psychology, however, occupational psychology has also been heavily influenced by theories originating from the management sciences and from the discipline of organization theory and behavior. It is possible to identify at least four main theoretical perspectives that have had a discernible impact upon occupational psychology: (1) individual differences, (2) applied social psychology, (3) applied cognitive psychology, and (4) management science. Table I summarizes these approaches and historic influences.

2.1. Individual Differences

The individual differences approach to personality traits, cognitive abilities, values and attitudes to work, and workplace behavior has a long and illustrious history in occupational psychology. Much of the seminal work in trait differences was undertaken using correlational and factor analytic statistical techniques (including the work of Eysenck, Cattell, and later by commercial personality test developers). The structure of human intelligence was similarly informed by early, seminal research mostly in laboratory experimental studies by scholars such as Galton and Vernon. Debate still continues as to whether tests of GMA or measures of specific abilities best predict subsequent job performance. Current meta-analytic research internationally converges on the finding that GMA tests are highly valid stand-alone predictors of job performance and that measures of specific abilities do not add significant incremental validity. Other individual differences have also been investigated, including attitudes toward work, leadership beliefs, work-related values, and interpersonal attitudes and style. All of these strands have tended to adopt a trait perspective, presuming that these

TABLE I
Underlying Theoretical Approaches to Occupational Psychology

	Individual differences	*Applied social psychology*	*Cognitive psychology*	*Management theory*
Basic tenets	Individuals differ along stable, measurable dimensions that affect observable behaviors	Interpersonal and group processes affect group productivity and individual well-being	Individual cognitions, information processing capacities, and decision heuristics affect individual task performance	Organizing processes and leader behavior are critical determinants of organization success
Primary level(s) of analysis	Intra-individual Inter-individual	Inter-individual Intra-group Inter-group	Intra-individual	Organizational Inter-group
Illustrative topics of influence	Selection	Group processess	Decision making	Organization structure
	Personality at work	Attribution theory	Human–machine interface	Authority relations
	Stress proneness	Innovation and creativity		Organization success
		Conflict		
Key models and constructs	Five Factor model (FFM) of personality	Group development models	Information overload	Contingency theory
	Type A, Type B	West's Four Factor model of team innovation	Heuristic biases	HRM and human capital
	General mental ability (GMA)		Information Integration Theory (IIT)	Organizational culture and climate

dimensions are stable, replicable over time, and causal of behavior across workplace situations, a perspective for which there has indeed been considerable research support.

2.2. Applied Social Psychology

As can been seen in Table I, the primary level of analysis for which applied social psychology research has been influential in occupational psychology is that of the workgroup or team in organizations. This has been a far more recent theoretical influence than individual differences research, however, and so there has been considerably less attention in the field paid to group-level issues than to individual-level issues. In more recent years, organizations internationally have moved more toward team-based structures of work organization, spurring considerable growth in research interest at this level of analysis. Much social psychological research into groups has been conducted in laboratory experimental conditions, leaving open to doubt whether their findings can be applied to real-life workplaces and business organizations. Recent occupational psychology research has been grounded upon earlier findings, models, and theories in applied social psychology but has examined phenomena in the workplace. Phenomena such as workgroup innovation, team performance and productivity, team decision-making among boards of directors, process losses in work teams, distributed cognition and decision-making, and conflict in workgroups have come to the fore. This research has irrevocably shifted the focus of occupational psychologists away from an exclusively individual level of analysis toward issues of interpersonal, group, and inter-group processes in organizations. Multilevel designs, that is, research measuring variables and effects across two or more levels of analysis (individual, workgroup, organizational), are becoming more common as a result.

2.3. Cognitive Psychology

Historic influences from cognitive psychology upon certain issues in occupational psychology are evident, as illustrated in Table I. Many of these influences concern individual-level issues of ergonomics, human–machine interface, and individual error in task functions. In certain countries (the United States and Germany, for instance), such applications were traditionally termed industrial psychology, with large-scale funding for often being provided by military research institutes or in industries where error consequences are of particular concern (e.g., nuclear power generation).

2.4. Management Theory

Distinct from other areas of applied psychology, occupational and I/O psychology has been heavily influenced by more macro-analytic theories, approaches, and issues addressed generally in the management sciences. Table 1 suggests some key examples, but the field has been impacted by research from several sub-disciplines in management, principally organization behavior (OB), organization theory, HRM, and organization development (OD). Most of these approaches have adopted an organizational level of analysis in considering performance and processes across entire organizations or multiple subunits as part of a work organization. Thus, management research in occupational psychology has been highly pragmatic in orientation, almost always involving quantitative field studies or even qualitative case studies, and has examined practical problems and processes as they occur in work organizations. In essence, this fourth approach has considered the managerial implications of occupational psychology through major topics such as leadership, HRM processes and organizational performance, theories and models of organizational structure, organizational change and development, culture and climate as determinants of success, and so forth.

3. KEY TOPICS IN OCCUPATIONAL PSYCHOLOGY

This section of the encyclopedia is composed of seven separate articles covering key topics in occupational psychology. They are as follows:

1. Contracts
2. Job stress and burnout
3. Occupation and gender
4. Personal initiative and innovation
5. Personnel selection
6. Safety and risks
7. Work teams

These purposely concentrate on the individual and team levels of analysis, given the historic background and theoretical approaches to occupational psychology

delineated above. An overview of each of these topics is provided in the following sections.

3.1. Contracts

Employment contracts and psychological contracts form the basis of the employer–employee relationship. The exchange nature of what each gives relative to what is expected to be received from the other is crucial, and can be termed the "mutuality" of expectations and obligations. While employment contracts are usually written, formal, and legally enforceable, psychological contracts are sets of mutual expectations that may be differentially construed by individual employees and managers in the workplace.

Psychological contracts especially are changeable in nature, open to violation by either party, and may be open to negotiation and renegotiation over time. The relative power of each party to negotiate a favorable contract is therefore important. Perceived violations of the psychological contract by employees have been found to be frequent, but research into perceived violations on the employer's side by underperforming employees is sparse.

3.2. Job Stress and Burnout

Stress and burnout at work represent major themes of research attention and concern within occupational and I/O psychology internationally. Stress is defined generally as the interaction between job environment stimuli and individual responses when the individual does not possess sufficient resources to cope, thus resulting in strain. So-called "interactional definitions" of stress emphasize the interaction between stimuli and individual responses as opposed to earlier, more limited, non-interactional conceptualizations of job stress. "Transactional definitions" imply that stress is a result of the transaction between the individual and his or her work environment.

Job stress and strain has been modeled in various ways, each guided by the definition of stress adopted by the researcher. Interactional and moderator models, which are more complex in their design and their inclusion of different types of variables, have become more prominent over recent years. A range of causal variables resulting in job stress has been identified, including work role demands, role conflict, role ambiguity, role overload, work complexity and multitasking, the home–work interface, the work–family balance, and financial and other non-work problems. Wider causes include business travel demands, relocation, expatriate working, bullying and mobbing at work, harassment, job insecurity, and redundancy. Increasing attention has thus been devoted to stress coping mechanisms and individual differences in coping with stressful events and work environments. Organizations, concerned by the potential costs of absence caused by stress and burnout, have invested in Employee Assistance Programs (EAPs), which are designed to support individual employees in a confidential and constructive manner. "Burnout" has been defined as "a special form of strain" characterized by feelings of emotional exhaustion, depersonalization, a lack of personal accomplishment, and a manifest inability to cope with even basic, ongoing demands in the work environment.

3.3. Occupation and Gender

Gender differences in career prospects, earnings, seniority of job roles achieved, and occupational choice have concerned occupational psychologists for many years. Women have been found to earn less than men, be concentrated into certain industries, and be more likely to work in certain lower authority job roles, and may have imposed upon them a "glass ceiling" of how high up the organization they can progress in terms of promotion to higher grade managerial job roles. A greater proportion of women work in part-time jobs and in service industry job roles, largely due, it has been suggested, to family commitments and traditional stereotypes of the mothering role and the housewife. These differences, still present in several developed economies cross-nationally today, have been fundamentally influenced by historic trends and precedents.

The socialization of boys and girls into potential work roles in organizations is also thought to have an impact. Gendered job roles, that is, traditionally male- or female-dominated occupations, may be held up as desirable during child development. For boys these may include military roles, emergency services, senior management, technical and scientific careers, and executive functions. For girls, in addition to family being emphasized, helping roles such as nursing, the caring professions, and service roles may be held up as stereotypical examples. Occupational values and choice may thus differ between the genders, and this spills over into differences in leadership role choice and style between males and females. Anti-discriminatory legislation in several countries that makes it unlawful

to discriminate on the grounds of gender does not appear to have eliminated these problems. Organizational efforts to enhance and manage diversity have shown positive outcomes, but many inequalities remain, and further research and practical intervention is called for.

3.4. Personal Initiative and Innovation

As organizations have moved away from job specialization and Taylorian scientific management as methods of work organization, the importance of creativity, innovation, and personal initiative in the workplace has grown. The need for product or service flexibility, adaptability to changing business environments, matrix- and team-based organization structures, and more flexible forms of working have all heightened the importance of work role creativity and team-level innovation. Creativity has been defined as the generation of new ideas, whereas innovation involves both idea generation and solution implementation in the work setting. Personal initiative is a related construct, involving self-starting, proactivity, and persistence, the outcome of which may be deemed to be innovative by observers.

Innovations develop over time not in a neat, linear fashion or through discrete step-by-step phases, but generally as iterative, complex cycles of development–regression–further development. Factors associated with innovativeness at the three interrelated levels of analysis in occupational psychology have been identified. At the individual level, these include personality, motivation, cognitive ability, knowledge, and job characteristics. At the work group level, these incorporate team dynamics and structure, climate, processes, and leadership style. Organizational level variables include structure, climate and culture, size, resources, work processes, and physical environment. Linear, curvilinear, threshold effect, and indirect relationships between these variables and innovation have been suggested or found. More recently, multilevel designs, possibly using more robust meta-analytic techniques, have been called for to clarify this mass of individual study findings.

Applied occupational psychology findings have countered a predominant but fallacious pro-innovation bias common in management texts that suggests that all innovations are positive and that the more innovation, the better for organizational performance. Rather, applied psychology research suggests that some types of innovations can be stressful for the initiator, problematic for the workgroup to assimilate, or conflict-provoking in the wider organization where other departments or subunits are affected.

3.5. Personnel Selection

Similar to stress and burnout, personnel selection has long been a major area of research and professional practice in occupational psychology. Based historically in individual differences theories and tenets, selection more recently has expanded its concentration of levels of analysis toward the team level and issues of person–work group fit. Applicant as well as organizational decision-making has also received considerably more research attention, but the traditional model of personnel selection is based upon fitting the person to the job through the administration of various prescreening and candidate assessment predictor methods.

Seminal usage and development of meta-analysis methods in selection have generated definitive research findings over the criterion-related validity of different predictors. The best stand-alone predictor of subsequent job performance across all complexities of jobs, and in different countries worldwide, is general mental ability. Other methods add to the prediction of the wider criterion space, including work sample tests, some dimensions of the "big five" personality framework, and integrity tests. Historically much maligned, interviews, particularly structured interviews, have been found in meta-analytical research to have far better predictive validities than previously thought. Other methods that demonstrate near-zero predictive validity include age, graphology, and other so-called alternative methods.

The dilemma that the most valid and reliable predictors are not necessarily the most popular with practicing recruiters continues to concern researchers in personnel selection. Other important research has investigated fairness issues (adverse impact or unfair discrimination) in selection in different countries. Driven by differing national legal frameworks, the stringent minority rights legislation in the United States has led to research there being far more developed than in other countries. The minority groups protected under legislation differ across countries, but usually depend upon the pattern and origins of immigrant workers to a host country.

Increasing research attention is being devoted to applicant reactions, so-called "banding" of minority

group scores to mitigate the effects of underprediction, how best to combine the information gleaned from different predictors at different stages in a selection process, expatriate selection, selection for team-working, and new technology in the recruitment process.

3.6. Safety and Risks

Fay and Tissington proposed that unsafe behavior and errors lead to higher risks and therefore may cause accidents to happen on the work floor. Unsafe behavior is defined as "behaviors that are not in accordance with safety rules and procedures." Error can be seen as "an intentional action which does not result in the attainment of the goal [whereby] prevention of the error must have been under the control of the person(s) [involved]." Research has adopted two different approaches toward safety and risks: the person-centerd approach, focusing on the role of individual difference variables, and the engineering control solutions approach, seeking to design safety into the work system by acknowledging human factor variables. Three subsequent levels of prevention are central to this approach: the design of an error-proof system, engineering safeguards, and finally, installation of warning systems. Interventions striving to enhance safety aim at changing characteristics of the workforce, changing the work itself, and/or altering elements of the work context.

3.7. Work Teams

A work team is defined by Kozlowski and Bell as "a collective of two or more individuals that performs organizationally relevant tasks, shares one or more common goals, interacts, exhibits task interdependencies, manages boundaries, and is embedded in a broader organizational context." As organizations have become flatter and delayered, team-working has become the predominant form of work organization in many post-industrial economies. In the past, Taylorism and the specialization of work dominated many large corporations' mode of work organization, but organizations have become increasingly dependent upon team-level co-operative efforts for their success. Team effectiveness models of an input–process–output nature guided classical research into workgroup effectiveness, yet have become increasingly complicated by researchers' attention to moderator and mediator variables and relationships.

Contingency models of team composition and structure emphasize the need to base team formation upon environmental and task demands, including such design issues as team size and diversity of members (i.e., homogeneity versus heterogeneity). Certain outcomes warrant a more diverse composition (e.g., idea generation, adaptability), whereas others suggest a more homogenous member composition (e.g., routine task performance, innovation implementation). Models of team development have also moved from a more prescriptive-descriptive stance to becoming more situation-dependent (i.e., contingent) as research has become more sophisticated. Fixed stage models have been supplanted by more recent punctuated equilibrium models.

Emerging research centers on cognitive and affective models of team motivation, the efficacy of team-building interventions, task interdependence and performance including knowledge structures, and highly pragmatically important issues of team creation, team-building, and performance improvement attempts. Little research appears to support the commercially driven consultancy market in this regard despite repeated calls by researchers for such evidence. Future research in occupational psychology into team-working in organizations is likely to reflect its increasing importance for organizational success, and the growth in research in this area is thus likely to continue apace.

See Also the Following Articles

Contracts ■ Industrial/Organizational Psychology, Overview ■ Job Stress ■ Job Stress and Burnout ■ Occupation and Gender ■ Personal Initiative and Innovation ■ Person–Environment Fit ■ Personnel Selection ■ Safety and Risks: Errors and Accidents in Different Occupations ■ Vocational Interests ■ Work Environments ■ Work Motivation ■ Work Role, Values Sought in the ■ Work Safety ■ Work Teams

Further Reading

Anderson, N., & Prutton, K. (1993). Occupational psychology in business: Strategic resource or purveyor of tests? *The Occupational Psychologist*, *20*, 3–10.

Anderson, N., Herriot, P., & Hodgkinson, G. P. (2001). The practitioner-researcher divide in Industrial, Work, and Organizational (IWO) psychology: Where are we now, and where do we go from here? *Journal of Occupational and Organizational Psychology*, *74*, 391–411.

Anderson, N., De Dreu, C. K. W., & Nijstad, B. A. (2004). The routinization of innovation research: A constructively

critical review of the state-of-the-science. *Journal of Organizational Behavior*, *25*, 147–173.

Arnold, J., Cooper, C. L., & Robertson, I. T. (1998). *Work Psychology: Understanding Human Behaviour in the Workplace*. London: Financial Times Publishing.

Arnold, J., Silvester, J., Patterson, F., Robertson, I. T., Cooper, C. L., & Burnes, B. (2004). *Work Psychology,* 4th ed. Harlow, UK: Pearson Education.

Cleveland, J. N., Murphy, K. R., & Stockdale, M. (2000). *Women and Men in Organizations*. Hillsdale, NJ: Lawrence Erlbaum.

Cooper, C. L., Dewe, P. J., & O'Driscoll, M. P. (2001). *Organizational Stress: A Review and Critique of Theory, Research, and Applications*. Thousand Oaks, CA: Sage.

Hodgkinson, G. P., Herriot, P., & Anderson, N. (2001). Re-aligning the stakeholders in management research: Lessons from industrial, work and organizational psychology. *British Journal of Management*, *12*, 41–48.

Klein, K. J., & Kozlowski, S. W. J. (Eds.). (2000). *Multilevel Theory, Research, and Methods in Organizations: Foundations, Extensions and New Directions*. San Francisco, CA: Jossey Bass.

Kozlowski, S. W. J., & Bell, B. S. (2003). Work groups and teams in organizations. In W. C. Borman, D. R. Ilgen, & R. Klimoski (Eds.), *Handbook of Psychology: Industrial and Organizational Psychology* (Vol. 12, pp. 333–375). Chichester: Wiley.

Landy, F. J., & Conte, J. M. (2004). *Work in the 21st Century: An Introduction to Industrial and Organizational Psychology*. New York: McGraw-Hill.

Myers, C. S. (1926). *Industrial Psychology in Great Britain*. London: Jonathon Cape.

Myers, C. S. (1936). Autobiography. In C. Murchison (Ed.), *A History of Psychology in Autobiography*. Worcester, MA: Clark University Press.

Porteous, M. (1997). *Occupational Psychology*. London: Prentice Hall.

Salgado, J. F., Anderson, N., Moscoso, S., Bertua, C., & De Fruyt, F. (2003). International validity generalization of GMA and cognitive abilities: A European meta-analysis. *Personnel Psychology*, *56*, 573–605.

Schalk, R., & Rousseau, D. M. (2001). Psychological contracts in employment. In N. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswesvaran (Eds.), *Handbook of Industrial, Work and Organizational Psychology* (Vol. 2, pp. 133–142). London: Sage.

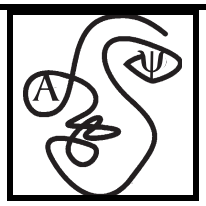
Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*, 262–274.

Shimmin, S., & Wallis, D., (1994). *Fifty Years of Occupational Psychology in Britain*. Leicester, UK: British Psychological Society.

Sinangil, H. K., & Avallone, F. (2001). Organizational development and change. In N. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswesvaran (Eds.), *Handbook of Industrial, Work and Organizational Psychology* (Vol. 2, pp. 332–345). London: Sage.

Vernon, P. E., & Parry, J. B. (1949). *Personnel Selection in the British Forces*. London: University Press.

West, M. A. (2004). *Effective Teamwork*, 2nd ed. Oxford: British Psychological Society/Blackwell.



Occupation and Gender

Eva Bamberg
University of Hamburg, Hamburg, Germany

1. The Gender-Specific Employment Market
2. History of I/O Psychology
3. Concepts and Results
4. Explanations
5. Career Advancement for Women and Managing Diversity
6. Practical and Scientific Perspectives
Further Reading

GLOSSARY

diversity management Planned steps to increase the diversity of the workforce and to integrate minorities in organizations.
gender Social construction, based on sex.
gender-related interventions Interventions designed to equalize men and women with regard to the workplace.
gender-specific division of labor Different types of work for men and women.
gender stereotypes Generalized beliefs about men and women.

The gender-specific division of labor in society implies that men and women practice different occupations, and consequently differ with regard to income, status, power, influence, and career perspectives. This gender-specific division of labor is supported and kept strong by processes of socialization and assignments. In the past few years, a number of interventions have been designed to advance the developmental potential of both sexes.

1. THE GENDER-SPECIFIC EMPLOYMENT MARKET

In most industrialized nations, the percentage of working women has risen continuously during the past few years, whereas the percentage of working men has remained almost the same. Approximately 80% of all men and 60% of all women practice an occupation. Although women and men share an equal basis of knowledge, acquired in school, career perspectives differentiate in horizontal as well as vertical aspects of occupation. The horizontal aspects of occupation concern the type of occupation and trade; the vertical aspects of occupation concern inner structural hierarchies. More men than women are suitably employed after educational training. Women more often practice work characterized by low demands and low decision latitude. Employment uncertainty affects women more than men; women more often work part-time.

Even when men and women hold similar jobs, men are provided with higher salaries. Female managers earn less than their male counterparts. Women in Europe earn approximately 15% less than men, and women in the United States earn approximately 25% less than men. These differences in salaries affect not only the sexes but also families. Female-headed families are one of the most economically vulnerable segments of the population.

Men and women differ with regard to their career perspectives. In Europe, for example, men are twice as likely to be in managerial positions. Only 3% of top

©2004 Elsevier Inc.
All rights reserved.

management positions are occupied by women. Worldwide, men dominate executive functions.

Gender-specific working conditions are found within the psychology community as well. Male psychologists are more likely to be provided with higher salaries, secured positions, and more executive functions than their female colleagues. Another phenomenon of the employment market can be observed for the psychology job market as well: When an organization or profession becomes more "feminine"—female employees increasingly predominate—the public value and influence of its members decrease considerably.

The gender-specific division of labor is not limited to paid work: Beginning in early childhood, girls perform more household chores than boys. Numerous empirical studies show that women do chores more than men. These findings are not limited to housewives and part-time working women; they also hold for full-time working women.

Work tasks are gender specific as well. Twice as many women as men do volunteer work in the social field, and nearly three times as many men volunteer in the area of politics. Men nevertheless mainly hold executive functions and prominent posts in the volunteer segment.

Work is divided in a gender-specific way. Type of occupation and gender are therefore important. Industrial/organizational (I/O) psychology focuses on premises, conditions, and consequences of work and division of labor within society. Gender and occupation should also be integrated.

2. HISTORY OF I/O PSYCHOLOGY

In the early years of I/O psychology, researchers reported sex-related characteristics of work and noted gender differences. The following is an example:

> Early Observations of Gender and Occupational Behavior
>
> "Sex differences can be directly identified methodically, based on any profession tasking both sexes in a parallel manner. Female and male tailors are differing widely concerning the manual aspects of their occupation. A man treats and processes fabrics even on a subjective point of view, applying strength, energy, and a dominance that should be called systematic. A tailoresse (female tailor) treats her fabrics somehow almost intuitional, playful, delicate, without any systematic, totally according to her own individual fittings." (p. 922)
>
> From Griese, F. (1928). *Psychologie der Arbeitshand.* Berlin: Urban & Schwarzenberg. Translated by E. B.

Differences between the sexes based on physiological aspects (e.g., muscular strength and operation of the cardiopulmonary system) and psychological aspects (e.g., intelligence, motivation, and commitment) were stated; specific sex-related aspects were noted.

The thesis of monotony resistance of women, developed in approximately 1920, stated that women engaged in repetitive work feel less bored and less monotony than men. Because repetitive work prevents struggling with constant decisions and allows daydreaming, women, who seek no satisfaction in their jobs, regard this kind of work to be convenient and desirable.

These statements are still made even today, but they have never been scientifically proven. A number of early studies by I/O psychologists integrated women (e.g., the Hawthorne study). However, no scientific discussion concerning the sexes of the test subjects and related consequences for the studies took place (and still do not take place). Other studies involved women only marginally in their samples. This has to be kept in mind especially regarding studies examining the consequences of working conditions and tasks on performance, health, and well-being. Until the 1980s, the issue of occupation and gender was almost totally ignored. Reacting to criticism about the neglect of women within industrial psychology research, a number of studies were performed that increasingly considered women and related results to gender-related topics. Kohn and Schooler, for instance, in their widely quoted study on occupation and personality, questioned men first and later their female partners, applying a longitudinal design.

Since the 1980s, an increasing number of studies have focused on women, resulting in increased research on gender-specific differences. Applying this new framework, research has focused on identifying and evaluating gender-specific differences (e.g., the degree of distinction in gender-specific values and interests and the degree of distinction in leadership behavior between men and women) and evaluating the differences in gender-related aspects within certain theoretical models (e.g., the connection between working conditions and personality).

3. CONCEPTS AND RESULTS

Common knowledge holds that gender-specific division of work is based on the differing work potential

of women and men. Applying this point of view, boys and men are mainly focused on career aspects, whereas women tend toward family aspects. Men are mainly interested in scientific/technical professions, which provide better career opportunities, whereas women focus on caring and nursing professions; men show materialistic interests, and women show social interests. Men are technically oriented and they impose their will; therefore, men succeed easier and are especially qualified for prominent positions of executive functions. Women are more cooperative. The following concept is regularly pointed out: Men and women differ according to the values related to the importance of work; therefore, consequences differ as well. For instance, for women unemployment is stated to be a minor problem because they have alternative roles.

How does common knowledge compare with scientific results? The following sections discuss three concepts of I/O psychology, focusing on occupation and gender.

3.1. Occupational Values and Goals

A number of studies have focused on whether occupational values and goals of men and women differ. Intrinsic values (social relations and job characteristics) and extrinsic values (salary, promotions, and occupational safety) are often separated. With regard to male and female frameworks of occupational values, there are only minor differences. Accordingly, intrinsic values are more important to women, and extrinsic values are more important to men.

The most prominent differences occur for young, married employees. The older the workers, the less prominent the differences become. This is due to the fact that men's focus on extrinsic values diminishes and their focus on intrinsic values increases with age.

A very striking connection was made by research examining the career orientation of students. The main sample showed only few differences between men and women. Focusing on the students' progression, the intrinsic career motivation of women is less prominent at the end of their scientific qualification period than at the beginning, whereas that of men is more prominent at the end of their scientific qualification period than at the beginning. Accordingly, there is a more prominent orientation on career aspects for men at the end of their scientific qualification period than for women.

Although the previously discussed results on occupational values show only few gender-related differences, these are found to be prominent by examining specific occupational plans and goals.

For many decades, research has elicited women's plans to reduce or discontinue occupational practice in favor of household matters. Correspondingly, especially women include housework in their future plans. This can be explained by the different social roles of men and women. Congruent to the gender-specific division of labor, men consider their future occupational practices to involve paid work. Their traditional responsibility is securing the family income. This injects special importance in career-related aspects, especially if entering the occupational market while simultaneously starting a family. Women view their working lives as involving paid employment as well as the family. Women have multiple roles. They are less likely to accept negative aspects or disadvantages of paid work than are men.

3.2. Leadership

Based on the previously mentioned gender-specific difference in employment within the management segment, leadership aspects have to be considered especially important. Here, the focus is on the occupational history and promotion of managers and on leadership behavior.

Studies have shown no significant difference between occupational values and motives of male and female managers. Compared to their male counterparts, female managers are more often single and have fewer offspring. Their occupational history can be compared to a "chimney," leaving space only for the paid work.

Mainly qualitative research indicates that men and women differ with regard to their identification of factors that enabled them to enter their career and gain promotion. Women refer more often to prejudices and systemic barriers—hindrances. Women in prominent positions with executive functions are more closely watched and they cannot use internal networks as effectively as men.

The prototype of the successful manager matches mostly male gender stereotypes and less often female gender stereotypes. The often validated phenomenon of think manager/think male is closely related to gender stereotypes, being the basis of judgments concerning leadership behavior. Despite major correspondences, these cause the difference in judgment of behavior of male and female managers. The position of female manager is considered as a token status. Women token workers experience a high degree

of visibility, isolation, and gender-stereotyped roles, causing discrimination.

Gender-related differences in leadership behavior have been the focus of many research projects in recent years. A more participative or personal-oriented style of leadership of women was noted especially by qualitative research. However, applied meta-analysis was unable to verify these findings. Research on leadership behavior only produced gender-related differences by applying laboratory-based methodologies. Field studies indicated only marginal gender-related differences. This could be due to the occupational socialization of the manager, corresponding with the development of specific skills. These skills build the central foundation of executive functions and are developed independently of being male or female.

3.3. Occupation and Health Issues

The average life expectancy of women is approximately 7 years longer than that of men. At first, this is not remarkable. Compared to women, men engage in behavior with significantly higher risks: Men smoke and consume alcohol more often than women, they abuse drugs more often, they less often have a healthy diet, they take advantage of health care programs less regularly, they are less sensitive toward signs of exhaustion, and they are involved in accidents at work more often. However, men are not more prone to sickness compared to women. Based on studies relying on self-reports, women suffer in a more intense way from psychological distress.

A number of explanations have been given for this finding. First is the more prominent emotional openness of women, which places women more at risk of psychological distress. This is more consistent with the female gender stereotype than the male, especially with regard to psychological distress and sickness. Another explanation is based on the differing vulnerability of men and women. The genetic, physiologic, and hormonal constitution of women is responsible for the inconsistency concerning stress, demonstrated by morbidity rates. The final explanation is the difference in working and living conditions between men and women, including stress related to paid work and housework.

The results provided by a meta-analysis on the significance of gender-related differences with regard to the relation of work and health are inconsistent with the second explanation presented previously. The second explanation states that there is a moderating function of gender within the relation of work and health. Due to the contrariness of the presented findings related to working conditions and health, determining the most plausible explanation seems to be impossible. Such a determination would require significantly more information about gender-specific working and living conditions. Information on stress and resources is needed to gain deeper insight into working and living conditions. Information is needed that is unlimited to aspects of the practiced occupational matters alone. There is a lack of adaptable measuring instruments capable of analyzing the overall stress impact and resources.

4. EXPLANATIONS

In the past, gender research was primarily based on biological models. This frames the previously mentioned statements about I/O psychology and the differing vulnerability of men and women. Biological models are based on the assumption that genetic, physical, and hormonal factors establish (gender-specific) behavior. Current approaches are based on the assumption that in addition to biological aspects (sex), social aspects (gender) also have to be considered. According to socialization models, the difference between men and women and gender identity is established throughout human socialization based on processes of learning, such as model learning, imitation, differential enhancement, and gaining cognitive concepts.

Social structural–cultural models of gender highlight the specific division of labor applied by certain institutions (e.g., schools, businesses, and the military) in order to maintain their specific function. In addition to division of labor, especially gender-related division of labor, a distribution of status and power takes place. This sets the foundation for social structures, norms, rules, and behavioral standards accepted and executed by the participants. Current concepts of occupation and gender integrate the previously mentioned approaches.

Biologically determined sex implies socially determined gender: Men and women perform different tasks and different types of activities. This has effects for all types of work—housework, freelance work, voluntary work, and paid work—forming a horizontal and vertical division in the employment market that affects not only types of activities but also power and status.

By working within specific areas of work, men and women adapt different attitudes, actions, competencies, and abilities. Gender-specific attitudes or competencies are found whenever the types of activities differ between men and women.

Gender-typical conditions, the types of activities of men and women, affect gender-specific differences and gender-related characteristics and processes. This influences the common knowledge base about men and women and their occupational attributes considered typically male or female.

Regarding psychology, research has focused on gender stereotypes. Gender stereotypes are structured assumptions creating a connection between personality traits and social categorizations—constructions of manliness/womanliness.

As shown by research on gender stereotypes, adaptive–instrumental attribution (e.g., achievement oriented, logically oriented, and independent) particularly describes men, and integrative–expressive attribution (e.g., sensitive, warm-hearted, and emotional) particularly describes women. Gender stereotypes include more than attributes; they also include role-consistent behavior, occupational praxes, etc. Gender stereotypes contain statements about the basic way men and women are (sex trait stereotypes) and how they have to be (sex role stereotypes).

The gender-specific division of labor affects the way in which men and women are described, and it affects gender stereotypes (see Three Schoolboy Essays). Research has shown that gender stereotypes are more prominent in people whose professions can be classified as traditionally male/female.

> Three Schoolboy Essays
>
> 1934: "I have often wondered what profession I would choose if I was a girl and was at the end of my school years. Since girls have not got so many options. Most of them become sales assistants. Since girls are frail other professions cannot be chosen for them" (p. 111).
>
> 1946: "If I was a girl, I would choose becoming a sales assistant for a bakery. I would like to do so because there would always be food for me in a bakery. Nowadays, girls are looking for apprenticeships at a grocery. The reason for this being, food can really be scarce" (p. 112).
>
> 1986: "Girls have fewer job perspectives than boys because there are mainly technical jobs left. Being a girl, I would take a job dealing with plants because I think that girls know more about plants. The job of a hairdresser is good for girls as well because they have more delicate hands than boys do. Myself being a girl, I would become a florist because boys are not tender enough to seed and care for plants" (p. 115).
>
> From Bamberg, E. (1996). *Wenn ich ein Junge wär*. Göttingen, Germany: Hogrefe-Verlag. Translated by E. B.

Gender-related division of labor affects not only the way men and women are described but also the way they describe themselves. A number of studies indicate that girls and women state they have expressive attributes, whereas boys and men state they have instrumental attributes. This can be mostly found in cases of similar experiences. In other words, whenever the job is considered typically male or female, generalized attributes of the job are adapted in the framework of the personal self. Research has shown that jobs considered typically female support the self-acquisition of female attributes, whereas jobs considered typically male support the self-acquisition of male attributes.

Based on the concepts presented here, a correlation between gender and occupation can be shown by applying two different perspectives. The individual or person-centered perspective focuses on acquisition of competencies, attitudes, and self-descriptions within socialization, including occupational socialization. Sex-related division of labor leads to the development of gender-specific attributes, which again support division of labor. Women who describe themselves as having expressive attributes will choose professions in which they can realize these expressive attributes.

The situational or context-related perspective focuses on valuation, furtherance, attachment, and hindrances set up by the actors and their frame of action. This again supports the gender-specific division of labor. Women working in typically male professions are often exposed to hindrances or systemic barriers that are not found in professions considered to be typically female.

5. CAREER ADVANCEMENT FOR WOMEN AND MANAGING DIVERSITY

In order to optimize manpower and minimize employee turnover, in the 1970s a number of businesses started to employ plans for career advancement for women. These plans employed a number of different starting points or target areas:

1. Ensuring personal safety: This includes guidelines to prevent discrimination and the creation of special safety areas for women, such as break rooms or reserved parking places.
2. Reduction of horizontal gender-specific segregation: This includes plans aimed at the facilitation of access to specific occupations, such as scientific/technical-oriented occupations.

3. Interventions focusing on changing vertical segregation: This includes mainly incentives aimed at increasing the percentage of female managers (e.g., mentoring).
4. Incentives aimed at increasing the compatibility between occupation and family life: This includes guidelines concerning work hours, maternity leave (the German *erziehungsurlaub* is an extended version of maternity leave for educational purposes), and implementation of part-time work.

As demonstrated by the latter target area, career advancement for women and the Equal Rights Amendment as well as family-support programs share overlapping components.

As mentioned previously, occupation and gender can be portrayed by two different perspectives: the individual or person-centered perspective and the condition-related perspective. Gender-related interventions are aimed at the implementation of individually oriented ways of development and at the creation of conditions necessary to advance development potential for both sexes. Therefore, both perspectives must be considered.

Nearly all interventions and measures for personal and organizational development can be implemented for gender-related interventions and measures as well. These include placing women in prominent positions, considering gender-related issues with regard to employment matters and recruiting, designing career advancement paths for women (e.g., offering training programs, mentoring models, and coaching), and implementing flexibility concerning workplace and working hours. Recently, the influence of gender-related measures on public opinion and ranking of organizations has been observed. Campaigns such as the family-friendly index, family audit, or the election of "house husband" of the year are part of this trend. In the United States, and in Europe as well, career advancement for women and the implementation of the Equal Rights Amendment are being replaced by the concept of managing diversity. The most prominent difference is that the concept of managing diversity does not just take women into consideration and does not focus only on forms of discrimination. The new approach centers on advancement of the developmental potential of all social groups within an organization (e.g., same-sex groups, age groups, and ethnic groups). Regarding multicultural organizations, employees are integrated not only structurally on all positional and hierarchical layers but also on the informal layer of networks. This decreases tension and discrimination based on employees belonging to minority groups. Staff politics is based on the premise of equal opportunity. Norms, criteria, procedures, and methods are tested with regard to discrimination and equality and changed accordingly if necessary.

6. PRACTICAL AND SCIENTIFIC PERSPECTIVES

The central concept of occupation and gender is based on the type of work. Men and women acquire attitudes, competencies, and abilities based on their work activities. The processing of gender stereotypes and gender-typical self-adjustment depends on the type of work. Consequently, there are implications for the areas of practice and science. Practical measures aimed at changing attitudes or increasing competencies have to begin with job design. Classical interventions within organizations and firms, such as job design or organizational development, did not focus on gender-related issues explicitly. This implies the nonobservance of women-specific working conditions. Most measures concerning job design, for instance, are based on jobs held by male employees and not on jobs mostly held by women.

Therefore, future approaches should not focus on equal rights politics and management diversity alone. The development of methods and measures designed to implement conditions considered typically male or female, with regard to job design and personal and organizational development, is needed. Gender mainstreaming is becoming an increasingly more relevant intervention within organizations.

It is necessary for concrete working conditions to be considered in research. This approach implies the existence of theoretical models that enable work to be defined in diverse fields. It further implies the existence of measuring instruments that enable work to be analyzed in different contexts. Both the theoretical models and the instruments of measurement do not exist. As mentioned previously, I/O psychology's focus on paid work (and on specific forms of paid work) indicates the limitation of applied models and instruments of measurement. Further developmental work has to be undertaken.

See Also the Following Articles

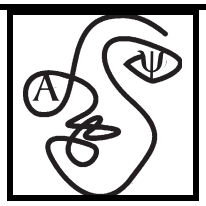
Gender and Culture ■ Gender and Education ■ Industrial/Organizational Psychology, Overview ■ Occupational Psychology, Overview

Further Reading

Cleveland, J. N., Murphy, U. R., & Stockdale, M. (2000). *Women and men in organizations*. Hillsdale, NJ: Lawrence Erlbaum.

Fredriksen-Goldsen, K. I., & Scharlach, A. E. (2001). *Families and work. New directions in the twenty-first century*. Oxford, UK: Oxford University Press.

Gilbert, L. A., & Rader, J. (2001). Current perspectives on women's adult roles: Work, family, and life. In R. K. Unger (Ed.), *Handbook of the psychology of women and gender*. New York: Wiley.



Openness to Experience

Robert R. McCrae
National Institutes of Health, Baltimore, Maryland, USA

1. Conceptualizations of Openness
2. Openness and Creativity
3. Openness and Vocational Behavior
Further Reading

GLOSSARY

absorption A trait characterized by intense concentration on an activity or experience.
divergent thinking The ability to generate many possible solutions instead of the single correct answer.
tolerance of ambiguity A willingness to deal with situations in which there is no clear-cut answer.
transformational leadership A leadership style required to make major changes in the direction or operation of an organization.

Openness to Experience is one of the dimensions of individual differences in the Five-Factor Model (FFM) of personality. Open individuals are characterized by an intrinsic interest in experience in a variety of areas and by a particularly fluid style of consciousness. Openness is strongly heritable and generally stable during adulthood. Open individuals are creative and adaptable, whereas closed people are pragmatic and down-to-earth. Recent research on Openness and vocational behavior shows that this factor is relevant to vocational interests, organizational change, and transformational leadership.

1. CONCEPTUALIZATIONS OF OPENNESS

Openness to Experience is the broadest of the five basic factors that comprise the Five-Factor Model (FFM), covering a wide range of loosely related traits. Among the relevant traits are tolerance of ambiguity, low dogmatism, need for variety, aesthetic sensitivity, absorption, unconventionality, intellectual curiosity, and intuition as measured by the Myers–Briggs Type Indicator. Perhaps because the links among such traits are not obvious, Openness is the least well understood of the five factors.

In studies of trait-descriptive adjectives, the fifth factor is usually called Intellect and is defined by terms such as perceptive, analytical, and reflective. Such words suggest intellectual ability, and Openness is sometimes confused with intelligence. In fact, intelligence (as measured by ability tests such as the Wechsler Adult Intelligence Scale) shows small correlations with Openness but forms a separate factor. Years of education is also modestly related to Openness; however, it is not clear whether education promotes Openness or whether open individuals pursue more education.

Perhaps the most widely used measure of this factor is the Openness scale of the Revised NEO Personality Inventory. In that instrument, Costa and McCrae provided subscales for six aspects of life to which individuals may be relatively more open or closed: fantasy, aesthetics, feelings, actions, ideas, and values. Thus, high scorers on the overall scale are imaginative,

©2004 Elsevier Inc.
All rights reserved.

responsive to art and beauty, attentive to their own feelings, willing to try new activities, intellectually curious, and unconventional. These traits appear to have both a motivational and a structural component. Motivationally, open people are receptive to new experience and to a deeper examination of their own thoughts, feelings, and values. Structurally, they have a permeable style of consciousness in which remote associations are easily made. In contrast, closed people are more comfortable with the world they know and tend to compartmentalize their ideas and feelings.

Although Openness appears to be an experiential style, it has pervasive social consequences. Open individuals tend to join liberal political parties and religious denominations. They seek friends who share their interests in the arts and tend to marry spouses who have similar open interests and values. Openness is also related to vocational interests and occupational behavior. Because Openness is so clearly expressed in attitudes and behaviors, there is generally strong agreement between self-reports and observer ratings on this factor.

In general, psychologists tend to be more open than closed, and many assume that Openness is a desirable trait. However, closed individuals prefer their own conservative and down-to-earth approach to life and would rate the desirability of Openness differently. Openness is associated with some negative outcomes, including nightmares, depression, and drug experimentation. Closedness is associated with authoritarianism and difficulty in adapting to change.

2. OPENNESS AND CREATIVITY

It is generally conceded that creativity is a complex phenomenon, depending on cognitive ability, training and experience, and the social context. However, creativity has also repeatedly been associated with personality traits and motives, and most of these (e.g., tolerance of ambiguity, preference for novelty, independence of judgment) can be considered aspects of Openness. Several features of open people predispose them to creative work. The fluid nature of their consciousness means that they can make unusual and remote (and sometimes creative) associations. The fact that they are low on the need for closure means that they can suspend judgment and seek alternative solutions instead of accepting the first that presents itself. The fact that they enjoy novelty means that they are motivated to seek new ways in which to accomplish tasks. Their love of beauty stimulates artistic creativity.

The aspect of intelligence thought to be most closely related to creativity is divergent thinking, assessed by tasks that require multiple answers to a single question. For example, a respondent may be asked to write as many sentences as possible with words beginning with a fixed set of letters or to list all of the possible uses for a brick. Tests of divergent thinking are moderately related to Openness, even after controlling for age, vocabulary, and years of education. Open individuals may be more interested in these tasks, or they may in fact possess this specific cognitive ability to a greater extent.

Openness has also been linked to creative performance in tasks such as writing stories, and it has been shown to predict real creative achievements. In 1999, Helson's pioneering studies of women's creativity over the life span demonstrated that those women who were to make substantial creative achievements had been characterized in college by traits related to Openness. Indeed, one easy way in which to understand the construct of Openness is by considering as case studies creative geniuses such as Jean-Jacques Rousseau and Leonardo da Vinci.

3. OPENNESS AND VOCATIONAL BEHAVIOR

During recent years, applied psychologists have begun to appreciate the importance of Openness for understanding vocational behavior. Openness is a major determinant of vocational interests. Open people prefer investigative and artistic occupations such as author, anthropologist, and research scientist, and they also endorse a wider range of interests. Closed people prefer conventional occupations such as banking and bookkeeping.

Open people welcome change, so it is not surprising that they are more likely than closed people to make mid-career shifts. Recent studies have also shown that they adapt more easily to organizational change. Part of that adaptability is tied to their willingness to seek information. Among new employees, more open individuals are more likely to seek out feedback about their performance from supervisors and coworkers, and they are more likely to regard learning a new job as a challenge. Meta-analyses have shown that open employees respond more favorably to training programs.

Because open men and women are creative and willing to seek new directions, they can become effective leaders of change, and studies have shown that Openness is associated with transformational leadership. Leaders, of course, also require other traits

(e.g., Conscientiousness) and skills, but for some leadership tasks, Openness may be a great asset.

It must be recalled that Openness is not inevitably superior to closedness. For many vocations and many roles within every organization, there are advantages to being conservative and conventional. Maintaining tradition can be as important as is creating innovation. Assessments of Openness can assist managers in assigning tasks and in understanding how employees are likely to perceive and react to organizational change.

See Also the Following Articles

Holland's Theory (Vocational Personality Types) ■ Leadership and Culture ■ Personality Assessment ■ Power, Authority, and Leadership ■ Psychometric Tests ■ Vocational Interests ■ Vocational Psychology, Overview

Further Reading

Helson, R. (1999). A longitudinal study of creative personality in women. *Creativity Research Journal*, *12*, 89–101.

Judge, T. A., & Bono, J. E. (2000). Five-Factor Model of personality and transformational leadership. *Journal of Applied Psychology*, *85*, 751–765.

McCrae, R. R. (1994). Openness to Experience: Expanding the boundaries of Factor V. *European Journal of Personality*, *8*, 251–272.

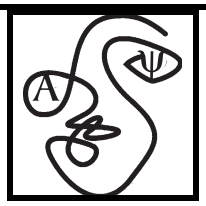
McCrae, R. R. (1996). Social consequences of experiential openness. *Psychological Bulletin*, *120*, 323–337.

McCrae, R. R., & Costa, P. T., Jr. (1997). Conceptions and correlates of Openness to Experience. In R. Hogan, J. A. Johnson, & S. R. Briggs (Eds.), *Handbook of personality psychology* (pp. 269–290). Orlando, FL: Academic Press.

Van Hiel, A., Kossowska, M., & Mervielde, I. (2000). The relationship between Openness to Experience and political ideology. *Personality and Individual Differences*, *28*, 741–751.

Wanberg, C. R., & Kammeyer-Mueller, J. D. (2000). Predictors and outcomes of proactivity in the socialization process. *Journal of Applied Psychology*, *85*, 373–385.

Wolfenstein, M., & Trull, T. J. (1998). Depression and Openness to Experience. *Journal of Personality Assessment*, *69*, 614–632.



Optimism

Christopher Peterson
University of Michigan, Ann Arbor, Michigan, USA

Nansook Park
University of Rhode Island, Kingston, Rhode Island, USA

1. Introduction
2. Contemporary Approaches to Optimism
3. Measures of Optimism
4. Outcomes
5. Development of Optimism
6. Deliberate Interventions
7. Conclusion
   Further Reading

GLOSSARY

agency Determination that goals can be achieved.
dispositional optimism Global expectation of optimism versus pessimism.
explanatory style Habitual style of explaining the causes of bad events, either with external, unstable, and specific causes (optimism) or with internal, stable, and global causes (pessimism).
hope Global belief that goals can be achieved (agency) by generating successful plans (pathways).
learned helplessness Passivity produced by learning that responses and outcomes are independent.
optimism Belief that the future holds more good events than bad events.
pathways Beliefs that successful plans can be generated to reach goals.

Optimism refers to the belief that the future holds more good events than bad events. Optimism is a popular topic of study among contemporary psychologists, who have approached it under the rubrics of dispositional optimism, explanatory style, and hope. Regardless of how optimism is measured, researchers usually find that it is associated with desirable outcomes, with one notable exception: Optimistic individuals underestimate likelihood of future risks. More is known about the development of pessimism than of optimism, although deliberate interventions can nurture optimism among youth as well as adults.

1. INTRODUCTION

Optimism represents a cognitive, emotional, and motivational stance toward the future. Thinking about the future, expecting that desired events and outcomes will occur, acting in ways believed to make them more likely, and feeling confident that these will ensue given appropriate efforts sustain good cheer in the here-and-now and galvanize goal-directed actions.

The term optimism is a relatively recent arrival on the historical scene, as is its ostensible cousin pessimism. In the 1700s, Leibniz characterized optimism as a mode of thinking, and Voltaire popularized the term in his 1759 novel *Candide*, which was highly critical of the shallowness of an optimistic perspective. Interestingly, pessimism appeared fully a century later, when it was independently introduced by Schopenhauer and Coleridge.

At least in their original forms, optimism and pessimism were not symmetric. Optimism as discussed by Leibniz was cognitive in its emphasis, reflecting a

©2004 Elsevier Inc.
All rights reserved.

reasoned judgment that good would predominate over evil, even if goodness were sometimes associated with suffering. In contrast, pessimism as discussed by Schopenhauer had an emotional meaning: The pessimistic individual was one for whom suffering would outweigh happiness. Someone can be optimistic in the cognitive Leibniz sense yet pessimistic in the emotional Schopenhauer sense.

As these notions evolved in the language, optimism and pessimism became juxtaposed, and the connotations of each spilled over into one another. The resulting composite is what psychologists usually mean by these terms, a dimension anchored by optimism at one end and pessimism at the other. Accordingly, optimism refers to a belief—perhaps wish would be a better term or even motive—that in the future good events and associated positive feelings will outweigh or be more likely than bad events and associated negative feelings.

2. CONTEMPORARY APPROACHES TO OPTIMISM

Contemporary writers approach this family of strengths in two ways. Some—usually social philosophers—treat these as features of general human nature to be praised or decried depending on the writer's own biases, whereas others—usually research-minded psychologists in the personality or clinical tradition—regard optimism and pessimism as individual differences, characteristics that people posses to varying degrees.

There are three well-known approaches within psychology to this strength as an individual difference. Each line of work has an associated self-report measure, each usually has focused on the consequences of the individual difference as opposed to the antecedents, and each has spawned a large empirical literature demonstrating that hope and optimism (or at least the absence of their opposites) are associated with desirable outcomes—positive mood and good morale; perseverance and effective problem solving; academic, athletic, military, occupational, and political success; popularity; good health; and even long life and freedom from trauma.

2.1. Dispositional Optimism

Charles Carver and Michael Scheier pioneered the study of a personality variable they identify as dispositional optimism, which is the global expectation that good things will be plentiful in the future and bad things scarce. Their overriding perspective is in terms of how people pursue goals, defined as desirable values. To them, virtually all realms of human activity can be cast in goal terms, and people's behavior entails the identification and adoption of goals and the regulation of actions vis-à-vis these goals. They therefore refer to their approach as a self-regulatory model. Optimism enters into self-regulation when people ask themselves about impediments to the achievement of the goals they have adopted. In the face of difficulties, do people nonetheless believe that goals will be achieved? If so, they are optimistic; if not, they are pessimistic. Optimism leads to continued efforts to attain the goal, whereas pessimism leads to giving up.

2.2. Explanatory Style

Martin Seligman and colleagues have approached optimism in terms of an individual's characteristic explanatory style. Those who explain bad events in a circumscribed way, with external, unstable, and specific causes, are described as optimistic, whereas those who favor internal, stable, and global causes are described as pessimistic. The notion of explanatory style emerged from the attributional reformulation of the learned helplessness model. Briefly, the original helplessness model proposed that following experience with uncontrollable aversive events, animals and people become helpless—passive and unresponsive—presumably because they have "learned" that there is no contingency between actions and outcomes. This learning is represented as a generalized expectancy that future outcomes will be unrelated to outcomes. It is this generalized expectation of response–outcome independence that produces later helplessness. However, the helplessness/explanatory style research tradition has infrequently examined expectations. Rather, researchers measure explanatory style and correlate it with outcomes thought to revolve around helplessness: depression, illness, and failure in academic, athletic, and vocational realms. In contrast to dispositional optimism, explanatory style is a construct concerning agency.

2.3. Hope

Rick Snyder's studies of hope in effect integrate these two visions of optimism—expectation and agency. Snyder traced the origins of his thinking to earlier work on expectations and goal achievement. According to Snyder's view, goal-directed expectations are composed of two separable components. The first is agency,

and it reflects someone's determination that goals can be achieved. The second is identified as pathways, which represent the individual's belief that successful plans can be generated to reach goals.

3. MEASURES OF OPTIMISM

Optimism (and/or pessimism) and hope (and/or hopelessness) have over the years inspired many measures. Most have been face-valid self-report questionnaires; other measures have relied on interviews, observer reports, or content analyses of written or spoken material.

Measures of optimism are usually based on what an individual says about what he or she usually believes. This strategy follows from the central role accorded expectations in the various definitions of these constructs; expectations are sensibly measured by self-report of beliefs. Each of the well-known contemporary research traditions has been facilitated by the existence of reliable and valid measures, such as the Life Orientation Test (LOT) to measure dispositional optimism, the Attributional Style Questionnaire to measure explanatory style, and the Hope Scale to measure agency and pathways. Each of these measures has been refined and revised several times to improve reliability and validity, and they are respected tools in the arsenal of research psychologists.

In the case of explanatory style and hope, there are separate self-report measures available for children and adults, as well as content analysis strategies for scoring these characteristics from written or spoken material. Ways to assess hope by observer report and from interviews also exist. Hope measures, and to some degree explanatory style measures, have spawned various domain-specific versions (e.g., hope concerning family life and explanatory style for academics). Dispositional optimism and pessimism have not been assessed among children and youth, although the straightforwardness of the LOT would presumably allow its use with early teens or even younger individuals.

A problem with respect to explanatory style is that questionnaire and content analysis measures do not impressively converge, although their correlates are usually identical. More generally, convergence among the different measures of optimism and hope has not been the subject of much research, although the typical study finds a modicum of agreement.

It matters whether measured expectations are about good events versus bad events. Therefore, the LOT has separate optimism and pessimism items, and these are somewhat independent and occasionally have different correlates. Explanatory style for good events is independent of explanatory style for bad events; the correlates are usually opposite for "good" versus "bad" explanatory styles and more robust for the "bad" style. This latter finding may be an artifact of the tendency in the explanatory style tradition to study negative outcomes (depression, disease, and failure).

4. OUTCOMES

As noted previously, optimism and hope predict many desirable outcomes: achievement in all sorts of domains (academic, athletic, military, political, and vocational), freedom from anxiety and depression, good social relationships, and physical well-being. More specific correlates include active problem solving and attention to problem-relevant sources of information, as well as conscientiousness, diligence, and the ability to delay gratification. People who are optimistic make "to do" lists, use day planners, and wear wristwatches; they also balance their checkbooks. All of these activities imply an active orientation to the future.

The only well-documented downside of optimism is its link with biases in risk perception. Optimistic people exaggerate the widespread tendency of most individuals to see themselves as below average for such dire occurrences as cancer and heart disease, failure and heartbreak, and so on. This bias can be a problem when reality matters and optimism leads one to neglect preventive or remediative actions. However, if people acknowledge the possibility of risk, the more optimistic among them take the most appropriate coping steps.

5. DEVELOPMENT OF OPTIMISM

Most of the relevant research has examined the environmental influences on explanatory style, and here it is clear what makes explanatory style more pessimistic: trauma and failure of all sorts. Less is known about what fosters an optimistic style, although an inconsistent literature finds that the explanatory styles of children and parents sometimes converge. Perhaps modeling is implicated, but so too might be shared family environment or even indirect genetic influence, as already discussed. If explanatory style is modeled, one needs to consider not just parents but also teachers, coaches, peers, and the media. Along these lines, a neglected but

likely important influence on the development of optimism is adherence to orthodox religion.

6. DELIBERATE INTERVENTIONS

The Penn Prevention Program (the Penn Resiliency Project) is a deliberate intervention to foster optimism. Schoolchildren are taught basic cognitive–behavioral strategies, especially disputation and decatastrophizing, for thinking about the causes of events in more optimistic fashion. This intervention makes explanatory style more optimistic and prevents the later development of depression and anxiety. It has also been demonstrated that both hope and explanatory style change in the course of individual psychotherapy with adults, especially of the cognitive ilk.

Missing in any of these intervention projects is an explicit focus on the fulfillments that presumably accompany optimism and hope. Thus, we know that the Penn Prevention Program makes emotional distress and even physical illness less likely, but it is unknown whether it makes youth happy and healthy.

7. CONCLUSION

Optimism and hope have had a checkered reputation since their introduction centuries ago. To call someone a Pollyanna is an insult, and to speak of cockeyed optimism or false hope is to underscore the foolishness of an endeavor that is based on an unrealistic appraisal of future risks. However, the overall benefits of optimism are not diminished by these occasional excesses. The most practical implication of optimism and hope research is to answer the question "Should one be optimistic?" by saying, "Yes, unless there is reason not to." In other words, optimism and hope should be regarded as cognitive strategies under the person's control—often useful but sometimes not—as opposed to immutable dispositions.

See Also the Following Articles

Personality Assessment

Further Reading

Buchanan, G. M., & Seligman, M. E. P. (Eds.). (1995). *Explanatory style*. Hillsdale, NJ: Lawrence Erlbaum.

Carver, C. S., & Scheier, M. F. (1990). Origins and functions of positive and negative affect: A control-process view. *Psychological Review*, *97*, 19–35.

Peterson, C. (2000). The future of optimism. *American Psychologist*, *55*, 44–55.

Peterson, C., Maier, S. F., & Seligman, M. E. P. (1993). *Learned helplessness: A theory for the age of personal control*. New York: Oxford University Press.

Seligman, M. E. P. (1975). *Helplessness: On depression, development, and death*. San Francisco: Freeman.

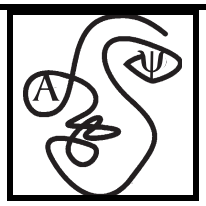
Seligman, M. E. P., Reivich, K., Jaycox, L., & Gillham, J. (1995). *The optimistic child*. Boston: Houghton Mifflin.

Sicinski, A. (1972). Optimism versus pessimism: Tentative concepts and their consequences for future research. *Polish Sociological Bulletin*, *25/26*, 47–62.

Snyder, C. R. (Ed.). (2000). *Handbook of hope: Theory, measures, and applications*. San Diego: Academic Press.

Tiger, L. (1979). *Optimism: The biology of hope*. New York: Simon & Schuster.

Weinstein, N. D. (1989). Optimistic biases about personal risks. *Science*, *246*, 1232–1233.



Organizational Culture and Climate

Mark F. Peterson
Florida Atlantic University, Boca Raton, Florida, USA

Ronald Fischer
Victoria University, Wellington, New Zealand

1. Basic Concepts
2. History of Organizational Climate Analysis
3. History of Organizational Culture Analysis
4. Organizational Climate Change and Consultancy
5. Organizational Culture Change and Consultancy
6. The Questionable Virtues of Strong Organizational Cultures
7. Are Organizational Climate and Organizational Culture Analogous to Characteristics of Larger Societies and Nations?
8. Institutional Analysis as an Alternative to Climate and Culture
9. Recent Literature Reviews and Compilations
Further Reading

GLOSSARY

institution In classical institutional theory, explicit structures (e.g., nations, organizations) created by legal means that give the institutions explicitly delineated authority to control their members and to enter into agreements with other parties; in neo-institutional theory, institutions are sets of understandings, actions, and interactions that characterize an identifiable, but not necessarily sharply or explicitly bounded, entity.

Likert scale A method for measuring attitudes, beliefs, or values in which a respondent answers a short question using a quantitative response scale, typically ranging from 1 to 5 or from 1 to 7, with each quantitative response linked to a verbal anchor reflecting a continuum (e.g., from *strongly agree* to *strongly disagree*).

organizational climate A term used to refer to how sets of individuals perceive and come to describe their work environment, their job, and/or organizational practices, procedures, and behaviors that are rewarded and supported, often in relation to the implicit goal priorities of the organization such as a climate for service or a climate for trust.

organizational culture A term used to refer to the configuration of values, implicit assumptions, and understandings that are expressed in overt statements, patterns of behavior, and interpretations of symbols that distinguish one organization from others.

strong organizational culture An organizational culture that is understood to have particularly strong effects on the ways in which organization members think and behave; it is usually contrasted with competing influences on organization members other than culture, including direct supervisory oversight, rules such as job descriptions and budgets, and explicit contracts.

values Desired states of affairs variously used to reflect implicit or explicit preferences, sometimes ordered into means–ends chains such that some values are viewed as preferred only as a means to achieving other outcomes, applied at various social levels from individuals across societies.

Organizational climate and organizational culture are terms that are used to comprehensively represent the social quality of organizations. The term "organizational climate" reflects psychological and sociological thinking about organizations, whereas the term "organizational culture" includes an anthropological element. Organizational climate is typically analyzed by using work attitude questionnaires. The scholarly literature

©2004 Elsevier Inc.
All rights reserved.

about climate often has been devoted to technical survey design and data analysis problems. These include how to reliably and validly measure individuals' attitudes and combine the answers to represent characteristics of whole organizations. Organizational culture is sometimes represented by surveys about values but is more typically analyzed by using interviews to gather stories or accounts of organizations from their members or by directly observing organizations either as an outsider or as a participant. Both terms are applied by managers and consultants who are interested in promoting or evaluating large-scale organization change.

1. BASIC CONCEPTS

Organizational climate and organizational culture both began as terms for comprehensive overviews of what an organization as a whole is like. The term "organizational culture" has retained its meaning as a comprehensive representation, whereas some uses of the term "organizational climate" have drifted from the original use. Organizational culture is most often used for referring to values and understandings, broadly shared in an organization or in some part of an organization, that influence the thoughts and actions of organization members. The term "climate" is now used for three purposes: (a) as psychological climate to describe questionnaire surveys that provide a broad profile of attitudes and perceptions of an organization's members covering more specific topics (e.g., job satisfaction, leadership, group dynamics), (b) as organizational climate to describe shared perceptions of organization members about their work environment and organizational policies and procedures that are likely to affect their well-being, and (c) as facet-specific climate to represent the distinctive value focus or profile of value priorities in an organization such as a climate for service or a climate for innovation. Although the concepts of organizational climate and organizational culture have come to be differentiated to some degree, each also has a heritage that reflects an interest in a holistic view of organizations.

2. HISTORY OF ORGANIZATIONAL CLIMATE ANALYSIS

The concept of organizational climate has a longer history than does the concept of organizational culture. The former derives from analyses during the 1930s of the qualities of a variety of social entities ranging in scope from a nation down to a group. Social climate was evoked to discuss the rise of fascism in Germany before World War II and the response of children's groups to adult leaders. The application of the social climate idea to organizations coincided with the development of the fields of organizational psychology and organizational behavior. Organizational climate was most often invoked when consultants and scholars working in these fields sought to provide a broad overview of the social side of an organization.

During the 1960s, climate surveys began to be used to ask organization members about what they experienced their organizational surroundings as being like. Employee responses were then combined to paint an overall picture of an organization that managers could use to evaluate the need for organization change and to plan and assess the effects of changes. Climate surveys included topics such as group dynamics, leadership, employee motivation, job characteristics, and management propensities to be supportive and fair. Many of these topics have become part of separate lines of research only loosely informed by climate research and are now most often encountered as distinctive literatures.

Likert, a social scientist and consultant with interests in attitude measurement, sociology, psychology, and organization change, inspired a great deal of the early organizational climate work. Most climate surveys use the now familiar Likert-type scale questions that ask people to rate work attitudes and perceptions. The typical scale for a climate item ranges from 5 to 7 points and includes verbal anchors ranging from "strongly agree" to "strongly disagree" (an agreement scale), from "to a very great extent" to "to a very small extent" (an extent scale), or from "very often" to "seldom or never" (a frequency scale). Climate surveys were one application to organizations of the questionnaire methods that began being designed for sociological and psychological research during the 1930s. Organizational climate surveys depended on emerging data analysis methods such as factor analysis during the early postwar years. The analysis of climate data became both more feasible and more complex as the field took advantage of developments in computer technology beginning in the 1970s. These methods made it increasingly practical to analyze data sets based on hundreds or thousands of respondents answering hundreds of questions.

The concept of organizational climate has evolved, in part, to sharpen its distinction from the concept of organizational culture. Three main perspectives have

been taken in the scholarly literature to distinguish it from organizational culture. First, climate has been conceptualized as individuals' representations or cognitive schema of their work environment. Individuals are thought to respond to situational variables in a manner that is psychologically meaningful to them. Therefore, this perspective focuses on the individual and is concerned with the psychological description of organizational conditions.

The second perspective focuses on shared perceptions of climate that can be used to describe either an entire organization or a work team. If individuals agree with each other about organizational policies, practices, or procedures that affect their well-being, these perceptions can be aggregated to the organizational level and used to describe the organization. Proponents of this approach have faced several methodological challenges. One is the difficulty of agreeing about appropriate ways in which to conceptualize and measure climate perceptions. Another is the establishment of statistical methods and cutoff criteria to determine that there is sufficient agreement or consensus among organizational members to believe that perceptions are sufficiently shared among organizational members to be used to describe the larger organization. This perspective is based on the argument that for shared perceptions to arise, it is necessary that (a) individuals interact at work, (b) individuals share a common goal or work toward an attainable outcome that predisposes them to collective action, and (c) individuals have sufficient task interdependence so that they develop a shared understanding of behavior. Because these conditions are more likely to be found in work teams than in the overall organization, proponents of this perspective argue that it is more likely that a shared climate develops at a group or team level rather than at an organizational level, especially if the organization is large, divisionalized, and multilayered. Consequently, this perspective on organizational climate has recently renewed the original social climate theme of group or team climates. Another offshoot of work treating climate as agreement or consensus has been the study of climate strength conceptualized as the level of agreement or consensus among organization members. Some argue that the effect of organizational climate on other variables, such as customer satisfaction, depends on how much members agree with each other, indicating a strong organizational climate.

Finally, climate has been used to represent a particular focus—a "climate for" something. Advocates of such facet-specific climates argue that using the term "climate" requires a specific referent such as a climate for service, a climate for change, or a climate for innovation. This perspective has revived the interest in climate research and has provided a new perspective for studying organizational phenomena with a strong applied focus. This facet-specific approach to climate also has been argued to sharpen the distinction of climate from the broader construct of organizational culture.

3. HISTORY OF ORGANIZATIONAL CULTURE ANALYSIS

Organizational culture is the newer of the two terms in organizational analysis, although (paradoxically) the use of the agricultural idea of culture as a social metaphor has quite a long history. Organizational culture is linked to analyses of societal culture having their origins during late 19th-century cultural anthropology. The basis of the metaphor is that just as societies engage in the culture of animals and plants, they also actively engage in social culture.

The relevance of anthropological perspectives on culture for managing organizations came to the worldwide attention of managers during the 1980s, although occasional references to the culture of organizations appeared several decades earlier. During the 1980s, Japanese management techniques such as quality circles were being widely imitated throughout the world, and Japanese companies were making many acquisitions throughout the United States, Europe, Asia, and Latin America. These business activities initiated debates about whether the success of Japanese organizations was due more to management practices that could readily be transferred or to distinctive qualities of Japanese culture that were either hard or impossible to imitate. Scholars were also showing increased dissatisfaction with limitations in climate research, notably with whether climate surveys were really more suited to representing the attitudes of individuals than to representing the attitudes of whole organizations. In addition, an even more basic set of debates was fomenting in academic management theory about how adequate quantitative techniques were for representing the social aspects of organizations. Although explicit discussion of Japan now has fallen out of use in the organizational culture literature, many characteristics of culture analysis can be traced to issues in Japan's effects on the business world during the 1980s.

The organizational culture literature has drawn attention to a number of aspects of organizations that had been overlooked previously. For example, in 1985, Schein drew directly from a functional cultural anthropology theory to suggest that all organizations need to deal with basic problems such as whether people are basically good, evil, or some combination of the two. The culture of an organization, notably its approach to control, was understood to be linked to the often unstated assumptions about such matters that tended to prevail. Other culture authors drew attention to the symbolism of objects such as the physical facility and the ways in which employees dressed. The anthropological tradition of interpreting stories as culture indicators was reflected in interest in organization stories and myths and in the distinctive roles that influential organization members played in transmitting these stories.

The organizational culture literature has also sought to use the methods developed for analyzing organizational climate and to adapt them to the new themes brought out by organizational culture analysts. In particular, climate surveys have focused on employees' attitudes and perceptions, whereas culture surveys have focused on values. This quantitative side of the organizational culture literature has encountered controversies about whether it is "really" culture analysis or is instead a form of climate analysis. Value surveys had in fact existed and been used frequently in psychology since the 1970s, but until the appearance of organizational culture, they had rarely been part of organization analysis.

4. ORGANIZATIONAL CLIMATE CHANGE AND CONSULTANCY

Organizational climate occasionally appears as a topic in an academic journal article, but it is more significant now as a routine part of organization assessments by corporate human resources managers in large organizations and by management consultants. Beginning in the 1960s, organizational climate surveys became part of the process of diagnosing organizations to identify how they need to change and providing feedback to promote change. For example, first a climate consultant would conduct a questionnaire survey for people throughout an organization. Afterward, the consultant would provide summaries of the survey results to various groups in the organization so that they could discuss what has produced the attitudes typical of their groups and what changes in the work system they might like to make. Alternatively, a consultant would provide managers with a summary report to inform their decisions about leadership training programs, improving the pay administration system, and/or implementing some other human resources change. Climate surveys are also used to assess the effects of either changes in an organization's competitive environment or new internal procedures initiated by management.

Following Likert, the types of organization improvements typically considered when climate surveys were first used were to increase employee participation in an organization's decision-making processes and to promote employees' satisfaction. These applications began during what is sometimes called the human relations period in organization analysis, that is, when management scholars and consultants typically had an egalitarian bias and accepted a view that employee satisfaction promotes employee performance. Although subsequent research has indicated that employee participation and satisfaction can contribute to organization performance under some circumstances, the consensus is that early advocates exaggerated their significance. More recent uses of climate surveys have retained the survey administration and feedback approach but have focused on identifying impediments that an organization is encountering toward achieving some particular goal such as service, trust, or innovation. Climate surveys remain one of the best ways in which to involve all organization members, regardless of their level, in diagnosing and assessing organizational changes.

5. ORGANIZATIONAL CULTURE CHANGE AND CONSULTANCY

The idea of organizational culture tended to come from consultancy-oriented approaches to organizations that appeared during the late 1970s and early 1980s and then were followed by scholarly assessments. Even more than organizational climate, organizational culture is now more firmly rooted in the management consultancy literature than in the academic management literature. One line of organizational culture consultancy advocates the virtues of establishing strong cultures. A second describes ways in which corporate leaders can show transformational leadership to effect cultural change. A third focuses more on diagnosing cultures using questionnaires in a way

that is reminiscent of organizational climate research. As detailed previously, organizational climate surveys focus on employee attitudes, beliefs, and perceptions about the people and groups around them, whereas organizational culture surveys focus on values. Some applications of culture surveys have been very similar to organizational climate applications, identifying values to the end of deciding whether and how to change them. Other applications have been to understanding and trying to promote effective work between groups that need to work together. These might be departments that traditionally have different values such as production and marketing, notably production efficiency versus customer responsiveness. They might also be parties in a merger or joint venture situation that are seeking to either interact more effectively or to anticipate and solve problems when combining into a single entity.

6. THE QUESTIONABLE VIRTUES OF STRONG ORGANIZATIONAL CULTURES

Many organizational culture consultants have recommended that organizations with strong cultures had advantages over those with weaker cultures. Specific views of what culture strength means vary, but most views have some common elements. These are that a strong culture includes understandings and assumptions that are broadly shared, as well as norms and values that are broadly accepted, throughout an organization.

Originally, there were two bases for advocating strong organizational cultures, both linked to presumed characteristics of Japanese society and the Japanese organizations that rose to world preeminence during the 1980s. One argument linked the advocacy of strong cultures to the idea of employee commitment. Theory and research related to organization commitment was developing at the same time as was culture research. It provided part of the explanation as to why Japanese culture promoted a generally more committed workforce than did most national cultures. Inspired by this belief, many scholars and consultants expected that a consistent link between commitment and performance could be documented despite the unsuccessful prior attempts to link satisfaction to performance. This hope was contradicted by research results that linked high commitment to low turnover but not to high levels of typically measured performance criteria.

A second argument was based on the idea of transaction cost economics that was being introduced to organization studies at about the same time. One category of transaction consists of those that occur between an organization and employees. Transaction cost thinking drew attention to the costs that organizations incurred in supervising people. One of the themes that linked anthropological understandings of culture to the economics of transaction costs was that cultural systems of norms and understandings could reduce the expense of monitoring employees. Japanese organizations were thought to economize on monitoring because employees knew how to behave appropriately due to a combination of general societal culture understandings and their socialization into an organization's culture. Some have argued that this understanding reduces the necessity of expensive ongoing performance evaluation and other monitoring systems. Based on both the expectation that employee commitment would have economic benefits and the view that the reduced need for monitoring would reduce monitoring costs, strong cultures came to be advocated. The virtue of strong organizational cultures continues to have its advocates, but the research base supporting the strong culture position has provided equivocal results.

7. ARE ORGANIZATIONAL CLIMATE AND ORGANIZATIONAL CULTURE ANALOGOUS TO CHARACTERISTICS OF LARGER SOCIETIES AND NATIONS?

The histories of climate and culture both are linked to analyses of larger societies. Culture theory was stimulated by questions about the business success of Japan, and climate research was influenced by attempts to understand how the social climate of Germany was transformed by fascism prior to World War II. The question of the relationship between nation characteristics and organization characteristics has continued. Part of the debate is about whether socialization into organization values has the same quality as does socialization into values while growing up as a child. Some would argue that the process is identical, whereas others would argue that organizations develop distinctive values either by selecting (or encouraging self-selection for) a particular type of value or by continually seeking to make salient a subset of those values developed during childhood.

8. INSTITUTIONAL ANALYSIS AS AN ALTERNATIVE TO CLIMATE AND CULTURE

The ideas of climate and culture have similarities to the idea of institution that is increasingly used in organization analysis. Institutions, like cultures, are viewed as patterns of understandings and actions that maintain stability in social systems and distinguish one system from another. Like organizational culture, institutions that appear as typical organizational practices, assumptions, and norms in one nation can be transferred to other nations, but the transfer requires adaptation to preexisting institutions. The term "culture" tends to be used in analyses that center on values, whereas institutions cover a broader range of phenomena that include established practices. The theory of institutions comes out of political science and is that field's way of dealing with the type of issues and phenomena that climate covers in sociology and psychology and that culture covers in anthropology.

9. RECENT LITERATURE REVIEWS AND COMPILATIONS

Numerous books offer managers guidance about particular ways in which to understand and change organizational climates and cultures. The seminal book *Organizational Climate and Culture,* edited by Schneider in 1990, was the first to focus on organizational climate and culture simultaneously and still marks a milestone in this area of research. Four more recent books provide scholarly reviews that expand and deepen these discussions and provide careful evaluations of many perspectives on climate and culture change. Trice and Beyer, in their 1993 book, thoroughly reviewed how the organizational culture concept developed in a way that highlights its anthropological roots and its distinctness from organizational climate. Academic articles by Denison in 1996 and by Verbeke and colleagues in 1998 discussed the historical development and conceptual overlap between the two constructs. Handbooks edited by Ashkanasy and colleagues in 2000 and by Cooper and colleagues in 2000 compiled an eclectic mix of perspectives on topics related to organizational climate and culture. These recent reviews and integrations suggest a number of directions being taken in organizational climate and culture research and consultancy in addition to those noted previously. In addition to covering many of the basic issues in climate and culture noted in the current review, some of the chapters in these handbooks linked organizational culture to other topics such as symbols, time, meaning, careers, individual attachment to organizations, and roles. In 2002, Martin provided a book-length review that frames the fields of climate and culture by using them as examples of struggles among competing basic conceptual perspectives on organization studies. Although organizational climate and culture are now less often the main focus of scholarly journal articles in management than they are in scholarly and applied books, some recent articles in the *Journal of Applied Psychology* and the *Journal of Organizational Behavior* show that academic interest continues. Furthermore, many topics that either first appeared or gained momentum through climate surveys and culture analyses are now studied separately.

See Also the Following Articles

Industrial/Organizational Psychology, Overview ■ Organization Development

Further Reading

Anderson, N. R., & West, M. A. (1998). Measuring climate for work group innovation: Development and validation of the team climate inventory. *Journal of Organizational Behavior*, *19*, 235–258.

Ashkanasy, N., Wilderom, C., & Peterson, M. F. (Eds.). (2000). *Handbook of organizational culture and climate*. Thousand Oaks, CA: Sage.

Cooper, C. L., Cartwright, S., & Earley, P. C. (Eds.). (2000). *The international handbook of organizational culture and climate*. New York: John Wiley.

Denison, D. R. (1996). What is the difference between organizational culture and organizational climate? A native's point of view on a decade of paradigm wars. *Academy of Management Review*, *21*, 619–654.

Hofstede, G. (2001). *Culture's consequences* (2nd ed.). Thousand Oaks, CA: Sage.

James, L. R., James, L. A., & Ashe, D. K. (1990). The meaning of organizations: The role of cognitions and values. In B. Schneider (Ed.), *Organizational climate and culture* (pp. 40–84). San Francisco: Jossey–Bass.

James, L. R., Joyce, W. F., & Slocum, J. W. (1988). Comment: Organizations do not cognize. *Academy of Management Review*, *13*, 129–132.

Likert, R. (1961). *New patterns in management*. New York: McGraw–Hill.

Lincoln, J. R., & Kalleberg, A. L. (1990). *Culture, control, and commitment*. Cambridge, UK: Cambridge University Press.

Martin, J. (2002). *Organizational culture: Mapping the terrain*. Thousand Oaks, CA: Sage.

Ouchi, W. G (1980). Markets, bureaucracies, and clans. *Administrative Science Quarterly*, *25*, 129–141.

Schein, E. (1985). *Organizational culture and leadership*. San Francisco: Jossey–Bass.

Schneider, B. (Ed.). (1990). *Organizational climate and culture*. San Francisco: Jossey–Bass.

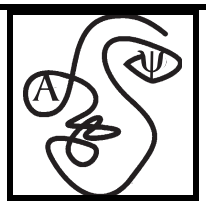
Schneider, B., Salvaggio, A. N., & Subirats, M. (2002). Climate strength: A new direction for climate research. *Journal of Applied Psychology*, *87*, 220–229.

Schneider, B., White, S. S., & Paul, M. C. (1998). Linking service climate and customer perceptions of service quality: Tests of a causal model. *Journal of Applied Psychology*, *83*, 150–163.

Scott, W. R. (1995). *Institutions and organizations*. Thousand Oaks, CA: Sage.

Trice, H. M., & Beyer, J. M. (1993). *The cultures of work organizations*. Englewood Cliffs, NJ: Prentice Hall.

Verbeke, W., Volgering, M., & Hessels, M. (1998). Exploring the conceptual expansion within the field of organizational behavior: Organizational climate and organizational culture. *Journal of Management Studies*, *35*, 303–329.



Organizational Diagnosis

André Büssing[†]
Technical University of Munich, Munich, Germany

1. The Concept of Organizational Diagnosis
2. Goals and Tasks
3. Methods
4. An Example: Organizational Diagnosis of Climate and Culture
5. Conclusion
Further Reading

GLOSSARY

climate and culture Important nonstructural organizational coordination mechanisms, representing an exemplary area for organizational diagnosis.

organizational diagnosis The systematic and scientifically based assessment, description, explanation, and prediction of regularly occurring experience, behavior, and performance of members of organizations and their interaction.

Organizational diagnosis is the systematic and scientifically based assessment, description, explanation, and prediction of regularly occurring experience, behavior, and performance of members of organizations and their interaction.

1. THE CONCEPT OF ORGANIZATIONAL DIAGNOSIS

Organizational diagnosis (OD) is a multistage process of assessment, analysis, description, explanation, and prediction. Features to be assessed are identified on the basis of theoretical considerations of relevant organizational characteristics and their relations (systematic). The diagnostic instruments have to comply with the standards of empirical research (scientific basis). The experience, behavior, and performance of organizational members, and also their interaction, are examined, setting aside individual differences in personality. The intention is not merely data gathering but also explanation and prediction.

Therefore, OD is connected to the understanding of organization and organizational psychology. Despite their diversity, organizational theories and models are based on the assumption that orders and rules exist in organizations and influence experience and behavior as well as performance of organizational members. The function of organizational psychology can be defined as the description, explanation, prediction, and modification of psychological phenomena in organizations. Its subject remains the organization with its multiple levels—individual and work task, group, unit, and organization—with the focus depending on the organizational theory being used. Consequently, the definition, goals, and tasks of OD are connected to organizational theories and models, and OD is always conducted from a certain position in organizational psychology.

There are several other definitions of OD; the majority date to the 1980s and do not originate from applied psychology but, rather, from management and organizational sociology. For example, Van de Ven and Ferry introduced a comprehensive approach to OD. Their framework of OD, which they called

[†]Deceased.

©2004 Elsevier Inc.
All rights reserved.

organizational assessment, draws on psychology, sociology, and management and incorporates a high degree of integration of levels, perspectives, and interdisciplinarity. Harrison and Shirom present a pragmatic approach to OD, which they define as investigations, using models and methods of the behavioral sciences, to examine the state of organizations, solve problems, or enhance organizational effectiveness. Whereas the term OD in this article is reserved for a predominantly psychological approach, and approaches from business administration, management, and organizational sociology are referred to as organizational analysis, Harrison and Shirom's approach is in between. Addressing all organizational levels, including individuals and jobs, with a focus on the application and implementation of OD, they represent a management-oriented position, primarily responsive to the needs of managers and consultants.

2. GOALS AND TASKS

OD represents a methodological framework for assessing, describing, explaining, and predicting experience, behavior, performance, and interaction of organizational members. This general goal of OD can be realized concretely in a variety of specific topics. The diagnosis of climate and culture and the diagnosis of quality in organizations are examples of the wide range of practical applications of OD. Consequently, the tasks of OD comprise a broad spectrum. Lawler *et al*. identified three groups of stakeholders interested in data and results of OD: members of the organization (employees and management), external stakeholders (e.g., government, shareholders, and responsible public bodies), and researchers.

Accordingly, they distinguish three kinds of tasks and applications for data and results of OD. For decision makers (management) in organizations, these comprise

- Organizational diagnostic data to improve human resource management
- Organizational diagnostic data to support staffing decisions about internal transfers and job changes
- Organizational diagnoses to prepare for and support organizational changes (i.e., conducting systematic and continuous analyses of deficiencies) pertinent to implementing strategic changes
- Organizational diagnoses to evaluate organizational program activities—structural (e.g., decentralization via team structures), psychological (e.g., quality circles), or clinical (e.g., addiction and rehabilitation programs)
- Organizational diagnoses to support decisions on the allocation of organizational resources (e.g., personnel, technology, and time)

The use of OD to provide decision makers outside the organization with important information is relatively rare. The following functions are the focus of OD for external decision makers:

- Organizational diagnostic data as the basis for investment decisions: There are no binding obligations for businesses or nonprofit organizations to provide external investors or controlling bodies—shareholders, governments, public advisory councils, or regulatory agencies—with decision-making information on the organization's state with regard to work and organizational psychological issues. However, findings of OD, such as information on qualification and knowledge, mobility, indicators of the quality of work life (e.g., satisfaction, commitment, and trust), can be highly relevant for decisions on technology, mergers, joint ventures, and strategic alliances.
- OD as a basis of a reporting system monitoring working conditions, health and safety, quality of work life, job attitudes, etc. to evaluate compliance with the criteria of humane work design: Similar to data on working accidents, absenteeism, or layoffs, these data could be considered in policy decisions regarding laws and regulations (e.g., tax, work, and social laws) and research programs. This application of OD will become more important as globalization and unification among countries (e.g., the European Union) increase.

The functions of OD for external stakeholders impose special requirements on the quality of data. This necessitates the authorization of independent external institutions, similar to auditing firms, to conduct OD because conflicting intraorganizational political interests make the validity of self-assessments questionable. Additionally, agreement on the applicable instruments is needed to allow for interorganizational comparability. These prerequisites illustrate that, to date, the application of OD in this context is more theory than practice.

The third domain for tasks and applications of OD is organizational research. OD can contribute to the development and testing of organizational theories. One criticism is the methodological one-sidedness of many organizational diagnostic approaches, primarily regarding measurement, which limits the comparability

of results and inhibits the advancement of organizational theory. To account for the diversity of organizational reality, Lawler *et al*. demand a renunciation of single methods in favor of multiple methods.

3. METHODS

OD incorporates methods of data gathering, data processing and analysis, formation of diagnostic judgment, and appraisal (i.e., the evaluation and presentation of results). In the following sections, the emphasis is on practical applications.

3.1. Instruments and Quality Criteria

Instruments from different disciplines, predominantly work and organizational psychology, economics, management, and organizational sociology, are used in OD. The choice of instruments largely depends on which aspects of experience, behavior, performance, and interaction of organizational members are relevant with regard to the focus and specific goals of the OD. Whereas, for example, for the diagnosis of organizational culture, the assessment of values and normative beliefs of organizational members is of particular relevance, for an OD focusing on quality, the perception and evaluation of internal and external customers are of vital importance. There are no comprehensive reviews on the instruments from the different disciplines; instruments are available from individual publications or books on certain groups of instruments. Table I lists seven ways of gathering data, for most of which a large number of instruments exists.

Frequently used are the analysis of documents and statistics; inquiries of organizational key persons, experts, and staff members; group discussions; and observations on the job of meetings, etc. Less common is the analysis of interactions in organizations by sociometry, network analysis, Bales' Interaction Process Analysis, or Kelly's relatively work-intensive Repertory Grid Technique. The choice of instruments from the variety of sometimes competing measures should be based on three aspects: focus (mainly if structures or processes are assessed), goal and purpose of the OD, and methodological strengths and weaknesses of the measures.

In practice, different methods and instruments are frequently combined, and often multiple methods are a response to specific strengths and weaknesses of the instruments. A step-by-step organizational diagnostic approach includes the use of nonstandardized and standardized methods in different stages. Initially, observations and interviews are conducted, mainly in unfamiliar, complex organizational areas. Next, processes and relationships are assessed primarily through standardized methods, such as questionnaires, guided observation, and archival analyses.

TABLE I
Data Sources in Organizational Diagnosis

Analysis of documents: organization charts, records, regulations, job instructions, job descriptions, organizational handbooks, special handbooks (e.g., for malfunctions and irregularities)
Organizational and economic statistics: general human resource data, absenteeism, working accidents, fluctuation, and production and investment data
Inquiries of key persons and experts: oral or written questioning of persons who are assumed to have implemented and/or control organizational regulations or who are involved in the implementation and/or control in certain subareas
Inquiries of staff members: oral or, more often, written questioning of organizational members on different levels
Observations on the job: meetings, etc.
Group discussions
Analysis of interactions: sociometry, network analysis, and interaction–process analysis

The quality criteria of psychological test theory represent a fundamental issue regarding the strengths and weaknesses of organizational diagnostic instruments. Because perfect measurement typically is not possible, different kinds of measurement errors, depending on the design of the instrument, are caused by insufficient objectivity (objectivity in use, analysis, and interpretation), reliability (estimated as test–retest, parallel, split-half reliability, or internal consistency), or validity (estimated as construct or criterion validity).

In addition to the essential criteria of objectivity, reliability, and validity and their necessary standards for conducting empirical research, practical and economic requirements have to be considered in the choice of instruments. Organizational diagnoses necessitate various kinds of investments. Demands for time and manpower, structure, semantic and visual comprehensibility of manuals and instructions, and simplicity of use and interpretation are important determinants of investments and have to be taken into account in the decision to use an instrument.

Despite methodological concerns, a goal conflict, which is inherent to an applied discipline such as OD and consistently has to be solved anew, remains between methodological–scientific demands and the pragmatic–technical requirements of practice. If organizational goals are achieved, despite a weak foundation of diagnostic measures and interventions, theoretical and methodological considerations sometimes may appear secondary.

3.2. Data Analysis

The analysis of data is an integral part of the organizational diagnostic process. Analytic techniques of empirical social research and applied statistics and specific techniques of individual psychometric diagnostics or process diagnostics are used. The former include descriptive statistics and statistical inference, such as analysis of variance to conduct comparisons of organizational groups, cluster analysis to generate diagnostic types, and discriminant analysis to generate and apply multivariate statistical selection strategies. For process diagnostic assessments of changes, analyses of variance with multiple measurement points or, in the case of many measurement points, more appropriately time series analyses can be conducted. Although a wide range of statistical methods exist for analyzing data from standardized OD, there is a lack of work on strategies for analyzing qualitative organizational diagnostic data, be it content analysis or interpretative methods. In this case, the diagnostician is referred to the literature on qualitative empirical social research.

3.3. Conducting Organizational Diagnosis

Conducting OD involves the phases of introduction, exploration, planning, main investigation, data processing and analysis, interpretation, summary, and presentation of results. In the introduction phase, diagnosticians develop a preliminary understanding of the relevant variables and dependencies, usually through talks and negotiations with clients, management, and employee representatives. Important variables are goals, potential benefits of results, available resources, and possibilities for participation of organizational members. Additional information on methods and anticipated problems of the OD is exchanged. A clarification of interests in this phase is essential for the success of the OD. An independent diagnostician has to take great care to avoid becoming biased toward one of the organizational groups (employee representatives and management). Both groups have to be completely informed during all stages. This also serves to prevent rumors and prejudices, which can result in misinterpretations of the OD as being part of a restructuring or downsizing program, which in turn can lead to low response rates, systematic biases, and the failure of measures.

In the exploratory phase, reconnaissance studies are conducted using largely nonstandardized interviewing methods, group discussions, and observations to consider and question issues and approaches and to prevent imbalances or restrictions in the diagnostic perspective. An evaluation of the sampling criteria for organizational units may also be indicated in this phase. Theoretical, methodological, and practical training may be necessary to align existing and required competencies and qualifications of investigators who are not experienced in conducting OD (e.g., freelancers and students).

The planning phase comprises the compilation, adaptation, and preliminary testing of the instruments, based on a model of the object of investigation, and considerations on their appropriateness regarding data processing and analysis. Additionally, various administrative issues of the investigation should be resolved in this phase (e.g., time, sampling, duration, location, information of participants, administration of instruments, and confidentiality of data).

In addition to these methodological, administrative, and concrete practical preparations, this phase can contain two higher level aspects of planning. First, cost–benefit analyses can be conducted for the purpose of the investigation and to help determine the diagnostic method (e.g., regarding requirements of time, personnel, materials, and sample size). Second, a meta-analysis of the scientific literature on methods and data collection can be performed if the reviews on certain issues of OD do not suffice for a reliable evaluation of the state of method development. The economy of the investigation should also be considered in the planning phase. Internet and corporate intranet offer good opportunities for its improvement. Currently, several sophisticated software packages for the social sciences facilitate online surveys and automated data processing (e.g., optical data scanning).

The planning phase takes on a central role in the OD. Not only does it set the course for the choice of methods but also a multitude of decisions are made during this phase that determine the organizational diagnostic process. The more accurate the planning,

the less fundamental problems will occur during the actual investigation. Despite careful planning, however, problems arise, at least in longer term organizational diagnostic projects. The four main problem areas involve the investigators conducting the OD, project controlling, changes in the examined organization, and unpredictable problems.

The lack of qualified personnel can cause staffing problems to become a crucial burden for any OD. Good leadership and human resource management are therefore imperative for success. Larger, longer term organizational diagnostic projects cannot be realized without regular feedback on compliance with schedule, budget, and quality standards. A systematic controlling of these aspects has to be incorporated into any project from the outset. Controlling time is usually less problematic than controlling costs and quality. This is caused, for instance, by the incalculable financial reserves needed to cover staffing gaps due to sickness or quitting of diagnosticians or the difficulties of controlling the quality of interviews, observations, etc. Changes in the diagnosed organization constitute a third problem area. Changes that interfere with the OD occur frequently in organizations but are difficult or impossible to predict (e.g., discontinuing of units due to market changes, changes in management, or unforeseen reactions of organizational members to the OD). Finally, there is always a residual category of unpredictable problems, ranging from unexpected method effects to the loss of data due to technical problems.

In the data-processing phase, compliance with predefined procedures for sampling and investigating and quality of data have to be assessed. This includes problems with methods (e.g., systematic nonresponse), data processing (e.g., errors in coding and screen forms for data input), or data management (e.g., confounding of variables and wrong recoding of items). Ensuring confidentiality of data is a major issue in the data-processing phase.

The actual phase of data analysis, carried out with statistical computer applications, should be preplanned, well specified, and hypotheses based, especially for larger data sets with numerous variables.

In the subsequent interpretation phase, results are evaluated with respect to the goal of the OD. Depending on the goal, results are used for inter- or intraorganizational comparisons (e.g., comparisons between units or divisions). Furthermore, results are interpreted with regard to hypotheses on relations (e.g., the perception of job insecurity is related to staff changes and a lack of cooperation/communication with the supervisor) and assumptions about causes and consequences of symptoms and interrelations (e.g., the perception of job insecurity results in decreased flexibility and increased absenteeism).

The OD concludes with the summary, report, and oral presentation of results for clients and interested organizational groups. Summary, report, and presentation significantly determine the organization's ability and willingness to accept and implement the results. Considering their importance for the success of the OD, these aspects often do not receive sufficient attention.

Active involvement of organizational members should be considered for every stage of OD, not just a subsequent phase of organizational design or development. This participation of affected individuals refers to a direct involvement in decision-making processes, including participation in the introduction, planning, and interpretation phases. Participation not only realizes aspects of democratizing and humanizing work life but also can represent a valuable support for the ecological validity of the OD. After all, who could give a better account of practices, processes, etc. than the persons involved, especially at the unit and work group level?

4. AN EXAMPLE: ORGANIZATIONAL DIAGNOSIS OF CLIMATE AND CULTURE

As nonstructural coordination mechanisms, organizational culture and organizational climate play a special role in organizational psychology. Both culture and climate are theoretical constructs that possess observable correlatives but are not directly observable. Observable cultural correlates are the use of language (communication culture), work and production means (technological culture), goals and mission statements (progress culture), rituals (traditional culture), safety standards (safety culture), etc. As a fundamental part of the organizational identity, organizational culture forms the character of the organization and its members both by attracting and assimilating members and by excluding or selecting them.

Although knowledge on the cultural and climatic situation of an organization attains strategic importance, culture and climate as constructs are not easily accessible for OD. Additional complications arise because a variety of models and definition of culture and climate in organizations exist and because culture and climate are not unequivocally distinguishable. Culture and

climate have in common the fact that they both refer to organizational characteristics that are based on communication and cooperation. However, whereas organizational culture refers to characteristics that are taken for granted, unquestioned, implicit, and customary (e.g., norms, symbols, and rituals), aspects of organizational climate are more consciously perceived, reflected, and judged in their connection with changing conditions.

For the diagnosis of organizational culture, standardized questionnaires assessing traits of organizations and organizational units (group and team) exist. According to Neuberger and Kompa, instruments for the diagnosis of organizational culture can be classified by method (quantitative vs qualitative) and by the relevant aspects of culture (sociocultural configuration vs latent factors). Consequently, four classes of instruments can be distinguished: interviews with key persons using, for example, questionnaires on leadership behavior (quantitative, sociocultural configuration); structured questionnaires assessing norms, values, etc. (quantitative, latent factors); structured questionnaires assessing stories, jokes, slogans, etc. on customs, rituals, and traditions in organizations (qualitative, sociocultural configuration); and spontaneous self-interpretations of meaning and function of cultural practices (qualitative, latent factors). This classification illustrates the different methodological approaches of OD for the example of organizational culture.

In 1990, Rousseau reviewed several quantitative measures for assessing organizational culture, including the Organizational Culture Inventory by Cooke and Lafferty, which is a widely used instrument.

The Organizational Culture Inventory assesses normative beliefs regarding the extent to which specific behaviors help members fit in and meet the expectations of the organization. Twelve cultural styles, each assessed by 10 self-report items using a 5-point Likert scale, are distinguished and arrayed in the underlying circumplex model by their orientation on task vs people and security vs satisfaction:

Cultural styles and sample items

1. Humanistic–helpful (e.g., helping others to grow and develop)
2. Affiliative (e.g., sharing feelings and thoughts)
3. Approval (e.g., "going along" with others)
4. Conventional (e.g., always following policies and practices)
5. Dependent (e.g., pleasing those in positions of authority)
6. Avoidance (e.g., waiting for others to act first)
7. Oppositional (e.g., pointing out flaws)
8. Power (e.g., motivating others any way necessary)
9. Competitive (e.g., turning the job into a contest)
10. Competence/perfectionistic (e.g., doing things perfectly)
11. Achievement (e.g., pursuing a standard of excellence)
12. Self-actualization (e.g., thinking in unique and independent ways)

The instrument has been applied in a variety of settings and cultural contexts. In a study of fund-raising organizations, it was shown that team-oriented cultural styles (achievement, self-expression, humanistic–helpful, and affiliative) were positively related to desirable attitudes (e.g., role clarity, satisfaction, and propensity to stay), whereas security-oriented beliefs (approval, conventional, dependent, avoidance, oppositional, power competitive, and perfectionistic) were negatively related to these attitudes and to organizational fund-raising success.

OD is not an end in itself, and this holds true for the diagnosis of culture and climate. The interest in reliable knowledge on culture and climate is reinforced by the expectation that a positive climate and culture support organizational goals, that their implications can be actively shaped, and that their benefits are significant. Attitudes, motivation, and values of organizational members are affected by culture and climate. For example, the relation of a positive communication culture and a climate of trust with high job satisfaction has been empirically confirmed.

5. CONCLUSION

Due to rapid sociotechnical changes in organizations, assessing the state before an innovation, controlling its implementation, and evaluating its success comprise an important function of OD. Changes in experience and behavior of individuals, due to technological innovations, impose demands on OD that are not consistently met.

The diversity of sociotechnical processes involves the extent to which different organizational levels are affected by a change in the system (from the macro-organization to the individual) and the nature and quality of changes (from changes in organizational structure and technological processes to changes in the possibilities for interaction and communication on the job). This diversity represents challenges for

OD and necessitates the development of integrative approaches, incorporating multilevel theory and methods, and broadening the perspectives adopted in the investigation of sociotechnical processes.

The considerable practical benefits of OD have not received sufficient attention. With the notable exception of the seminal work of Van de Ven and Ferry, examples of methodologically sound and comprehensive OD projects are rare. This, however, may change if the practitioners in organizations and consulting firms become familiar with the possibilities and the cost–benefit relation of OD. To achieve this, the comprehensiveness of OD and its necessity and position in the process of organizational design and development have to be emphasized and illustrated by examples. The misperception of OD as an interesting methodological approach that is dispensable for the solution of time-sensitive practical problems can be easily disproven with concrete practical results.

In addition to the various general and specific benefits derived from information and knowledge on the state of organizations, an important argument for the value of a psychologically based OD for businesses, managers, and consultants is the financial aspect. Similar to other areas with regard to work and organizations (e.g., worker safety and health protection), cost reductions attributable to OD are hardly calculable in detail. The profitability of OD cannot be appropriately evaluated by the traditional economic calculation of the costs of isolated measures but, rather, includes the optimization and integration of the whole system with regard to functionality, safety, etc. Only a comprehensive economic resource accounting, including a human capital assessment and using data from controlling and comparisons of costs and benefits, has the potential to generate a sustainable effect by constantly providing feedback on the results of OD and thus stimulating and reinforcing the motivation of decision makers, managers, consultants, and other involved groups (e.g., work councils). A comprehensive economic evaluation can be provided by extended accounting systems, which incorporate both monetary and nonmonetary goals (i.e., goals of the company and goals of the employees) and capital value and human capital reflected, for example, in the level of qualification and motivation.

Acknowledgments

This article is adapted from Büssing (2003). I thank Severin Hornung for help preparing this article.

See Also the Following Articles

Industrial/Organizational Psychology across Cultures ■ Industrial/Organizational Psychology, Overview ■ Organizational Culture and Climate ■ Organizational Structure

Further Reading

Anastasi, A., & Urbina, S. (1996). *Psychological testing*. Englewood Cliffs, NJ: Prentice Hall.

Büssing, A. (2003). Organisationsdiagnose [Organizational diagnosis]. In H. Schuler (Ed.), *Lehrbuch der Organisationspsychologie* [Textbook of organizational psychology] (3rd ed., pp. 445–480). Bern: Huber.

Cook, J. D., Hepworth, S. J., Wall, T. D., & Warr, P. B. (1993). *The experience of work. A compendium and review of 249 measures and their use*. London: Academic Press.

Cooke, R. A., & Lafferty, J. C. (1989). *Organizational culture inventory*. Plymouth, MI: Human Synergistics.

Cooke, R. A., & Rousseau, D. M. (1988). Behavioral norms and expectations: A quantitative approach to the assessment of culture. *Group and Organizational Studies*, *13*, 245–273.

DiPofi, J. A. (2002). Organizational diagnostics: Integrating qualitative and quantitative methodology. *Journal of Organizational Change Management*, *15*, 156–168.

Harrison, M. I., & Shirom, A. (1999). *Organizational diagnosis and assessment: Bridging theory and practice*. Thousand Oaks, CA: Sage.

Hays, W. L. (1994). *Statistics*. Cambridge, MA: Wadsworth.

Howard, A., *et al*. (Ed.). (1994). *Diagnosing for organizational change: Methods and models*. New York: Guilford.

Klein, K. J., & Kozlowski, S. W. J. (Eds.). (2000). *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions*. San Francisco: Jossey-Bass.

Lawler, E. E., Nadler, D. A., & Cammann, C. (1980). Uses of organizational assessment data. In E. E. Lawler, D. A. Nadler, & C. Cammann (Eds.), *Organizational assessment: Perspectives on the measurement of organizational behavior and the quality of work life* (pp. 25–118). New York: Wiley.

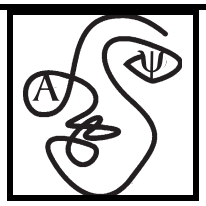
Miles, M. B., & Huberman, M. A. (1994). *Qualitative data analysis—An expanded sourcebook*. London: Sage.

Neuberger, O., & Kompa, A. (1993). *Wir, die Firma. Der Kult um die Unternehmens-kultur* [We, the company. The cult about the organizational culture]. Munich: Heyne.

Rousseau, D. M. (1990). Assessing organizational culture: The case for multiple methods. In B. Schneider (Ed.), *Frontiers of industrial and organizational psychology* (Vol. 3, pp. 153–192). San Francisco: Jossey-Bass.

Stevens, J. (2002). *Applied multivariate statistics for the social sciences*. Hillsdale, NJ: Erlbaum.

Van de Ven, A. H., & Ferry, D. L. (1980). *Measuring and assessing organizations*. New York: Wiley.



Organizational Justice

Paul E. Levy and Christina Norris-Watts
University of Akron, Akron, Ohio, USA

1. What Is Organizational Justice?
2. History of Justice in Organizations
3. Theoretical Conceptualizations of Justice
4. Antecedents of Organizational Justice
5. Consequences of Organizational Justice
6. Conclusion
Further Reading

GLOSSARY

distributive justice The perception that what one receives in terms of outcomes is fair.

fairness heuristic theory An approach to making justice judgments that relies on cognitive shortcuts as guides in decision-making behavior.

fairness theory An approach to making justice judgments that considers whether other alternatives could and should have been used as well as the extent to which they would have led to more favorable outcomes (an extension of the earlier referent cognitions theory).

instrumental model A perspective arguing that justice is important to individuals because the voice provided them allows them some control over outcomes.

interactional justice The perception that the interpersonal treatment that is received from one's coworker, supervisor, etc. is fair.

organizational justice Individual's perceptions of the general level of fairness in a specific organizational context.

procedural justice The perception that the procedures or processes used in a particular situation such as the workplace are fair.

relational model A perspective arguing that justice is important to individuals because being treated fairly suggests that one is an important individual to others.

Organizational justice refers to perceptions of fairness in an organizational context. Justice research focuses on theoretical conceptualizations of justice, how justice decisions are made, and the dimensionality of justice, as well as its antecedents and consequences. Fair treatment is highly valued in the 21st century workplace, where justice perceptions tend to impact on employee attitudes and behaviors as well as the effectiveness of organizations.

1. WHAT IS ORGANIZATIONAL JUSTICE?

Organizational justice is the study of people's perceptions of fairness in organizational contexts. The term was coined by Jerry Greenberg in the 1980s to describe individuals' interest in and concern with fairness-related activities taking place in various organizations such as one's workplace. Organizational justice can be viewed as a class of motivated behavior that is engaged in by different individual and contextual characteristics. Perceptions of organizational justice then result in cognitive and affective responses that guide ensuing behaviors. It did not take long for the concept to capture the attention of both researchers and practitioners, and this interest has resulted in numerous books,

©2004 Elsevier Inc.
All rights reserved.

journal articles, and conferences that have been devoted to organizational justice. In addition, when consultants talk to managers, employees, and CEOs about the "hot" issues confronting their organizations and the people of those organizations, the consultants are deluged with examples, anecdotes, and stories that revolve around organizational justice and its impact on the organization and its members.

2. HISTORY OF JUSTICE IN ORGANIZATIONS

Adams' work on equity theory is the foundation for psychological work on fairness. The focus of this work was on the fairness of outcomes, or what has become known as distributive justice. Organizational researchers recognized the potential effect of distributive justice on organizational functioning and, thus, it has been a major focus of organization-based research since the 1980s. However, as it became clearer that employees' reactions to organizational events were not based solely on their perceived notions of outcome fairness or distributive justice, the direction of research followed from work in social psychology and process-related issues began to be examined. Thus, procedural justice, or the fairness of the processes used to determine outcomes, became a major research focus and this has continued into the 21st century. Procedural justice is enhanced by providing employees with opportunities for voice or input into the process and by following fair and consistent rules. In fact, Leventhal suggested a set of rules to follow to ensure procedural justice that included accuracy, ethicality, and representativeness.

In the mid- to late 1980s, Bies suggested a third dimension of organizational justice that he called interactional justice. Interactional justice is believed to be related to procedural justice, but in particular it describes how decisions are communicated to employees and focuses on interpersonal treatment. That is, whereas procedural justice focuses on the fairness of organizational procedures involved in decisions, interactional justice deals more with how these decisions are conveyed to employees.

There continues to be disagreements among researchers as to the number of independent dimensions of justice. Some believe only in the traditional distinctions between procedural and distributive, where interactional justice is subsumed under procedural justice. Others maintain interactional justice as a third independent dimension of justice. It has also been suggested that interactional justice is composed of two dimensions: (i) interpersonal justice, which focuses on treating individuals with respect, dignity, and courtesy, and (ii) informational justice, which focuses on the explanations provided to individuals about the decision-making process.

3. THEORETICAL CONCEPTUALIZATIONS OF JUSTICE

Justice research has been driven by two main questions. First, why is justice important? Research has investigated the extent to which justice is critical to employees and why they value justice. Second, how do employees decide what is just? Researchers have developed theoretical models for examining and understanding this judgment process.

3.1. The Importance of Justice

Research has demonstrated that individuals value justice and behave, in part, based on their perceptions of fair treatment. One approach explaining why individuals value justice takes an instrumental perspective, arguing that individuals desire control over events and situations. Individuals perceive that they have control when they are provided the opportunity to voice their opinions, beliefs, and ideas. Ultimately, this instrumental model suggests that voice is used to exercise some control over situations, increasing the likelihood of favorable outcomes. A slightly different perspective is the relational model (formerly called the group-value model), which argues that individuals desire favorable relationships with others. Being treated fairly provides individuals with evidence or proof that they are valued by others, increasing their own sense of self-worth.

A recent perspective takes the position that justice is its own reward. The moral virtues model argues that justice is important to individuals based on an innate sense of right or wrong. In other words, individuals have a fundamental respect for the worth of human beings and, therefore, expect fair behaviors that are consistent with this respect. Finally, Cropanzano and colleagues attempted to integrate these various perspectives into what they call the multiple needs model, in which basic desires for control, belonging, positive self-regard, and meaning motivate a desire for justice.

3.2. Deciding What Is Just

Two major theories have been developed that purport to explain how individuals make justice decisions. Applied to organizations, these theories describe two different processes through which employees might evaluate work situations and draw conclusions about the fairness of treatment received. Referent cognitions theory, developed by Folger in the 1980s, posits that situations are perceived to be unfair when individuals can think of more favorable outcomes that *would* have occurred from alternative procedures that *could* have been used. If individuals cannot generate alternative procedures that would have led to more favorable outcomes, the situation is perceived as fair. In Folger's revision of this theory, fairness theory, a third requirement was added. Along with the *would* and *could* elements, he proposed that situations are only perceived as unfair if an alternative procedure *should* have been used according to a moral imperative. In sum, injustice is perceived when (i) unfavorable outcomes are experienced, (ii) alternative procedures could have been used, (iii) the alternative procedures would have led to more favorable outcomes, and (iv) these alternative procedures should have been used based on some ethical rationale.

Lind developed another perspective on the justice decision-making process in the 1990s that he termed fairness heuristic theory. In general, heuristics are shortcuts used to make decisions that free up cognitive resources for other purposes. With respect to justice decisions, Lind argues that individuals make fairness judgments and then use those judgments to regulate behavior. Research has shown that individuals' behaviors are more driven by early fairness information than later fairness information. In other words, once individuals make a fairness judgment based on available information, that judgment demonstrates stable properties and guides ensuing behavior.

4. ANTECEDENTS OF ORGANIZATIONAL JUSTICE

Now that justice has been defined and the prominent theories reviewed, it is necessary to consider variables that influence justice perceptions. These antecedents can be classified into three groups: individual differences, situational characteristics, and supervisory factors.

4.1. Individual Differences

Individual differences such as demographic variables, affectivity, and self-esteem have been investigated as potential antecedents of organizational justice perceptions. First, demographic variables such as age, gender, race, education, and organizational tenure seem largely unrelated to justice perceptions. This suggests that no particular category of individuals are more inclined to perceive injustice than any other category. Second, studies have found negative affectivity to be related to justice perceptions. Negative affectivity can be thought of as a personality disposition descriptive of individuals who have a generally pessimistic worldview. The findings suggest that those who have more negative affectivity tend to perceive more injustice in their organizations. In particular, this relationship is stronger for procedural and interactional justice than it is for distributive justice. Third, employees who have high self-esteem, manifested by self-confidence and positive perceptions of self-worth, tend to perceive more procedural justice in their organizations than do those who have low self-esteem.

4.2. Situational Characteristics

Organizational policies, procedures, and practices have been examined as predictors of justice perceptions. For instance, strong relationships have been demonstrated between the extent to which employees are provided with opportunities for voice and perceptions of procedural and distributive justice. This is consistent with the previous discussion about why justice is important to employees: Having input leads to perceptions of control, which is valued by the employee and results in impressions of justice. The degree to which organizations provide support and high-quality communication to their employees has also been linked to justice perceptions. For instance, employees who perceive more organizational support also view their organizations as being more procedurally just. Finally, one of the most consistent findings in the literature is that individuals react more strongly to unfavorable outcomes if those outcomes are accompanied by unfair procedures. In other words, fair procedures can partially overcome the negative effects that are usually associated with unfavorable outcomes. Individuals who experience unfavorable outcomes but who view the procedures as fair react more positively than do those who receive similar outcomes and view the procedures as unfair.

4.3. Supervisory Factors

The social context plays an intricate part in organizational functioning and it has been examined in terms of justice perceptions as well. In particular, leadership variables have been extensively studied with respect to justice perceptions. The relationship between leaders and subordinates has been the focus of a great deal of research in industrial/organizational psychology. This relationship is often operationalized as an exchange between leaders and subordinates. Leader–member exchange (LMX) theory has been developed to describe this social interaction. Research has found strong relationships between leader–member exchange and justice. In particular, more favorable leader–subordinate relationships are associated with higher perceptions of justice. Interestingly, consistent with the previous discussion of the interpersonal nature of interactional justice, the relationship between LMX and justice is strongest when justice is operationalized as interactional in nature.

In addition, transformational leadership—a leadership style in which the leader uses consideration and charisma to inspire employees—has been positively linked to procedural justice perceptions. Furthermore, transactional leadership—a leadership style in which the relationship between leader and supervisor is based on simple economic exchanges (i.e., meet these goals and receive this reward)—has been positively linked to distributive justice perceptions, as the theories would predict.

5. CONSEQUENCES OF ORGANIZATIONAL JUSTICE

One way to distinguish among the many consequences of organizational justice is to categorize them as attitudinal or behavioral in nature. When applied to organizational justice, this results in a variety of outcomes in each category.

5.1. Attitudes

One of the most studied variables in organizational psychology is job satisfaction or the emotional state resulting from the cognitive evaluation of what one's job provides to the employee in relation to one's expectations. It has been extensively studied with respect to justice, and strong consistent relationships have been demonstrated. The three major dimensions of justice have been consistently linked, in a positive way, to job satisfaction such that employees who experience organizational justice also tend to report greater job satisfaction than those who do not experience organizational injustice. Furthermore, satisfaction with one's pay level has been strongly related to procedural justice, but as predicted by the theory, it has been even more strongly related to distributive justice. The most well-established conceptualization of organizational commitment is termed affective commitment, defined as the psychological attachment to the organization based on common goals and objectives. Research has found that the three dimensions of organizational justice are uniformly related to levels of affective commitment. Employees who perceive more fair experiences at work also tend to be more committed to the organization. Justice has also been found to be related in predictable ways to other attitudes, such as trust in the supervisor, trust in the organization, supervisor satisfaction, and management satisfaction.

Other attitudinal variables of interest focus on employee reactions to various organizational systems. First, considerable research has demonstrated that reactions to the performance appraisal process are affected by perceptions of organizational justice. Individuals perceiving the performance appraisal process as fair tend to be more satisfied with the appraisal system and session, view the appraisal as more accurate, view the appraisal as more useful, and are more motivated to use it for developmental purposes. Second, considerable research has investigated the relationship between justice perceptions about the employee selection system and reactions to that system. These studies have found that both rejected and selected applicants respond more favorably to selection systems that they perceive as fair and job relevant. These favorable applicant reactions include the tendency to recommend the company to other potential employees, satisfaction with the application experience, and work self-efficacy.

5.2. Behaviors

Research suggests that justice perceptions predict employee performance. In particular, procedural justice perceptions have been strongly related to work performance such that employees experiencing fair work procedures are better performers than those experiencing unfair work procedures. Performance has been expanded to include contextual performance, which consists of those behaviors focused on maintaining the social core of the work environment rather than

those task-related behaviors that make up the technical core of one's job. In other words, contextual performance maintains and supports the work environment but is not part of the main tasks of the job. Research has found consistent, significant relationships between justice perceptions and contextual performance, such as altruism, good sportsmanship, loyalty, and courtesy, but these relationships tend to be small.

Counterproductive behaviors in organizations include employee deviance, such as stealing, sabotage, and both physical and psychological aggression. In general, both distributive and procedural injustice have been shown to predict these behaviors, but interactional justice has yet to receive much research attention with regard to counterproductive behaviors. Employees experiencing unjust outcomes or procedures have been more apt to exhibit counterproductive behaviors than have those experiencing fair outcomes or procedures.

6. CONCLUSION

A plethora of organizational justice research has been conducted, mostly in the past 15 years. Much of what goes on in organizations is affected by perceptions of organizational justice. Employees and supervisors alike make judgments about what is fair and unfair in their organizations. These judgments, in part, affect individuals' attitudes and behaviors and the interactions among individuals. Organizational functioning is indirectly affected via the processes involved in justice judgments and the ensuing behaviors. Be it in performance appraisal, selection, compensation systems, organizational change interventions, or the legal realities of organizational life, justice perceptions will continue to play an important role.

See Also the Following Articles

Leadership and Culture ■ Power, Authority, and Leadership

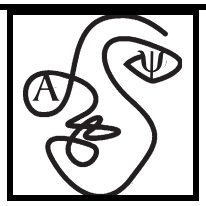
Further Reading

Cohen-Charash, Y., & Spector, P. E. (2001). The role of justice in organizations: A meta-analysis. *Organizational Behavior and Human Decision Processes*, *86*(2), 278–321.

Colquitt, J., Conlon, D. E., Porter, C. O. L. H., & Ng, K. Y. (2001). Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology*, *86*(3), 425–445.

Greenberg, J., & Cropanzano, R. (Eds.). (2001). *Advances in organizational justice*. Stanford, CA: Stanford University Press.

Subich, L. M., & Levy, P. E. (Eds.). (2001). Organizational justice: Implications for theory, research, and practice [Special Issue]. *Journal of Vocational Behavior*, *58*, 163–278.



Organizational Participation

Frank Heller

Tavistock Institute, London, United Kingdom

1. Introduction
2. Conceptual Schemes
3. Background
4. Examination of the Evidence
5. Conclusion
   Further Reading

GLOSSARY

codetermination Describes the German system, started by law in 1951 in the coal, iron, and steel industry, that gives workers 50% representation at the boardroom level. This was later extended, with lower worker representation, to other industries. Variations of this boardroom representation have been adopted in other European countries.

human resource management (HRM) A broader assessment of the role of employees than personnel management. The term indicates that employees are a resource (rather than a cost) in the way finance and capital are resources.

quality circle Japanese method of delegating quality responsibility to groups. It usually includes an intensive process of training. It has been adapted in various ways in other countries.

participative dialogue Derived from ideas developed by the German philosopher Jurgen Habermas and used particularly in Sweden to involve as many people as possible in organizational problem-solving processes.

ringo seido Japanese bottom-up decision making. The "ring" process starts at the lowest appropriate level and gradually moves up, with each person putting his or her "seal" on the paper. This results in joint consensual responsibility. The process is slow, but implementation is quick.

semiautonomy Derived from research in the Tavistock Institute, it represents a delegation of a degree of influence and responsibility to an individual or group for a defined task. The definition of the prefix "semi" is critical.

Organizational participation (OP) is about sharing influence and power in work organizations. This area of research has become important since World War II and has been called a number of names, such as employee involvement, high involvement management, participative decision making, partnership, joint consultation, industrial democracy, codetermination, and empowerment. Participation is still the most frequently used term and is often treated as a continuum. For instance, the research program at the University of Michigan used the terms system 1 (exploitative authoritative), system 2 (benevolent authoritative), system 3 (consultative), and system 4 (participative). In later research, particularly in Europe, the continuum was extended to include semiautonomy and delegation or complete control. In this extension, the term influence–power–continuum is sometimes used based on the assumption that participation can lead to degrees of autonomy or control and delegation of power in specified circumstances. There is a link with leadership and the decision-making process from the lowest level of organization to the boardroom and beyond to the national legal framework within which OP operates.

©2004 Elsevier Inc.
All rights reserved.

There is also a relationship between OP and the literature on industrial relations and organizational change. Collective bargaining, although part of the influence and power-sharing approach, is usually excluded from OP, as are cooperatives and share ownership; this article adheres to this tradition.

1. INTRODUCTION

Organizational participation (OP) can be treated from the micro perspective of the individual, the meso level of the organization, and the macro level of national policy. There are substantial differences between countries. The United States and Britain have voluntary schemes based on companies and sometimes on sections or divisions of companies. Most continental European countries have a legal underpinning for OP. For instance, the German codetermination scheme gives employees up to 50% representation at the supervisory board of their two-tier board structure. There are also elected representatives on works councils with legal rights to be consulted in defined circumstances. Similar legislation exists in most continental and Scandinavian countries.

However, formal representative participation does not require legislation. In Britain and the United States, such schemes are set up voluntarily by many companies sometimes related to collective bargaining. Formal schemes, sometimes called joint consultation or employee involvement, become part of company policy and frequently have a written constitution and elected representatives. Informal participation is quite different and can take many forms. It may be encourages by top management, but it is not usually accepted boardroom policy. It is normally left to individual managers, departments, or divisions and can vary from ad hoc meetings to regular forums. Informal participation is an aspect of managerial leadership and changes as managers leave or retire. A particular department may acquire a tradition of OP, which then lasts beyond a particular manager's style of operating. In voluntary formal and informal OP, the content of deliberation and the amount of influence available to different groups vary considerably.

In Japan, tradition and culture, rather than legislation, have produced formally sanctioned schemes of bottom-up management (*ringi seido*), the extensive use of quality circles, and the custom of lifelong employment in large companies. Some of these Japanese OP and human resources schemes have begun to weaken.

2. CONCEPTUAL SCHEMES

The goal of OP, whether formal or informal, is to achieve some organizational benefit, such as a more content workforce, more flexible work arrangements, less resistance to change, and increased productivity. Two theoretical models have emerged. First, the human relations model postulates a causal relationship between participation, work satisfaction and lower resistance to change, and increased organizational performance (productivity) (Fig. 1). As discussed later, evidence for this model is mixed. This model also lends itself to inauthenticity or pseudo-participation, often mentioned in the literature, since it may produce "a feeling of participation" through job satisfaction, without any real influence.

The second model is based on evidence that the positive outcome of OP seems to be due to the way in which participation allows existing experience and skill to be utilized. It is called the human resources model (Fig. 2). The assumed causality is participation → improved human resource utilization → improved organizational performance. Performance can take different forms, such as an increased propensity to innovate. The model does not exclude job satisfaction, although the direction of causality is not clear. The importance of this model is its emphasis on the utilization of competence, without which there can be no improvement of output. It follows that OP is not appropriate when employees do not have relevant experience and skill. It also virtually excludes inauthentic or pseudo-participation because the utilization of competence cannot be manipulated. The model suggests that any improvement in organizational performance is due to the utilization of existing experience and skill rather than to changes in attitudes.

Both models can be used in cross-sectional research with a single organizational level or department and therefore isolated from the total organization. Alternatively, they can be conceptualized as part of a systemic relationship within an integrated human

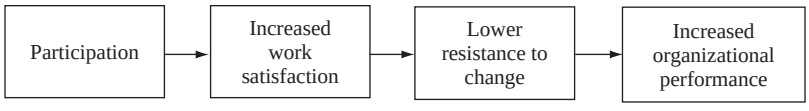


FIGURE 1 The human relations model. Possible feedback loops are not indicated.

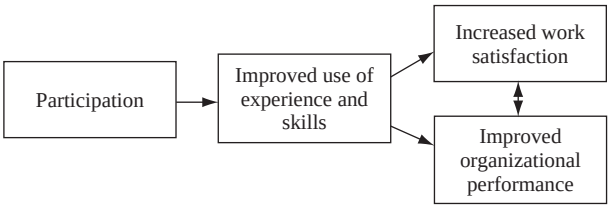


FIGURE 2 The human resources model. Possible feedback loops are not indicated.

resources policy. Furthermore, it may be necessary to recognize the impact of the external environment within an open systems framework. The effectiveness of OP may be influenced by new technologies, external uncertainties, national policy, legal provisions, and macroeconomic events. An open systems framework may therefore be appropriate (Fig. 3).

3. BACKGROUND

As economic competition increased in the decades after World War II, OP was frequently advocated as a way to improve organizational efficiency. This justification for OP has continued. In the 1980s, when Japan had a very competitive economy, European and U.S. managers began using Japanese methods such as quality circles and similar employee involvement techniques. Toward the end of the 20th century, economic conditions changed and Japan started to dismantle some of its human resource policy, particularly lifetime employment security. Nevertheless, during these competitive changes in Japan and the rest of the world, some aspect of OP has remained an organizational objective in most countries.

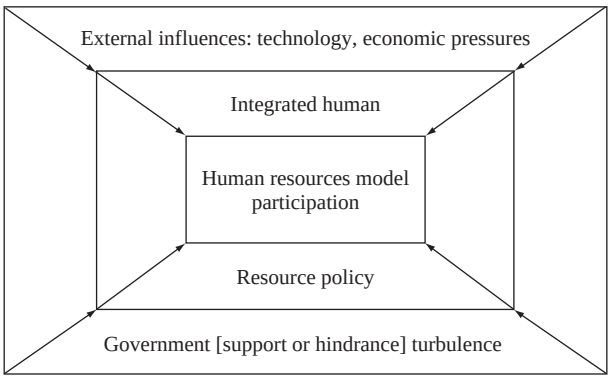


FIGURE 3 Open-system human resources model.

4. EXAMINATION OF THE EVIDENCE

4.1. Traditional Fieldwork

A considerable amount of the research literature concentrates on the lower levels of organization using a variety of quantitative methods and statistical designs. Much of this work, including meta-analyses, has been reviewed, for instance, by Locke and Schweiger, Cotton, and Strauss. The results do not present a uniform picture. OP frequently correlates with job satisfaction but much less often with improvements in organizational performance. Wagner reviewed 11 U.S. studies and concluded that even when there is a reliable statistical relationship between OP and performance, the correlations are so small that it is doubtful they have any practical significance and will tempt managers to adopt them on economic grounds alone. Other research, particularly if based on consultancy experience, such as that by Pfeffer, has found much more positive results; nevertheless, Pfeffer and others have been disappointed and surprised by the low uptake of OP work designs.

4.2. Cross-National Studies and Surveys

A number of large-scale cross-national studies on participative practices have been performed, especially in the United States and most European countries. Several studies cover the whole organization and use judgments from more than one level. Participative practices vary substantially between countries. One of the major reasons for this is the legal framework in which they are embedded (Fig. 3). When aspects of OP are mandatory, de facto influence sharing reaches the lowest levels of an organization.

The other consistent finding from large-scale national and cross-national studies is that the amount of distributed influence is very small. In a 12-country Industrial Democracy in Europe study, one of the scales asked people to describe their influence on a 5-point continuum from no influence (1) to very much influence (5). With regard to 14 decisions, the average low-level employee scored 2 (little influence). In relation to tactical and long-term decisions, their influence was much lower than 2.

4.3. Various Adaptations

The organizational function called personnel management has transmuted into human resource management

(HRM), which, as the name implies, gives greater weight to the role of human capital without diminishing the importance of economic and physical capital. Giving employees more voice plays a part in various HRM developments; for instance, semiautonomous work groups or teamwork have become popular and are often credited with improved efficiency as well as greater job satisfaction. As with other schemes, there are varying degrees of semiautonomy within teams. Sometimes, the power given to teams is minimal. In other cases, the team is able to reassign work sequences, layout, and task priorities and may be responsible for quality. Occasionally, more extensive power and responsibility are devolved to teams, such as choosing their own leader, setting targets and devising or distributing incentive schemes, agreeing on holiday rosters, job rotation, and job design, and communicating with management at a higher level. When teamwork is based on trust between management and employees and the group is assigned many of the tasks described previously, this is characteristic of a holistic system, which is discussed in the next section.

Quality circles have not always been successfully adapted to Western companies. They require skilled members, motivation, and incentives. In Japan, they were embedded in a nonadversarial industrial relations culture and could often call on the advice of outside experts.

The terms used to describe various HRM policies change frequently. Partnership is a term used to emphasize formal or informal influence-sharing arrangements between constituencies inside an organization or between an organization and an outside body (e.g., a private company and a public body). The development of virtual problem-solving organizations with people geographically dispersed, sometimes globally, has given rise to a form of interpersonal participation that can be said to create social capital.

As organizations and the environment in which they operate become increasingly complex, more emphasis is placed on a flexible workforce and the need for constant updating and learning. Among the characteristics required for an effective learning organization is a high level of motivation supported by a dispersal of influence to facilitate the acquisition of experience and skill. Equally important, but less frequently recognized, is the role of participation in reducing the underutilization of existing competence.

4.4. Systems-Oriented Research

Systems-oriented organization theory has been advocated since at least the 1960s through the writing of Katz and Kahn and the Tavistock team, among others. However, until recently, it has had little impact on HRM policy. Most research on OP is designed and assessed within one part of an organization. It might be favored by a particular manager or consultant and does not survive when the manager leaves the organization. Even if OP applies to a whole organization, it is usually isolated vis-à-vis other HRM schemes.

Given the inconclusive results from traditional isolated projects, based on the human relations model (Fig. 1), more holistic OP designs have received attention. They are more like a human resources model operating within a systems perspective (Figs. 2 and 3). Several facets of an organization are clearly interdependent and are treated as a system, such as accounting. Managers at all levels work within the accepted set of interrelated accounting procedures. There is evidence that a systems design of HRM that includes OP can yield positive results. There are packages of HRM schemes, sometimes called high-involvement management (HIM), that include participation and a bundle of measures, such as teamwork, compensation–incentives, training, quality circles, broad job description, and similar work practices. Research suggests that such interrelated HIM–OP measures create synergy, leading to increased effectiveness, efficiency, and/or productivity.

Another relevant experience derives from Norway and Sweden. A number of individual sociotechnically designed OP projects were started in the 1960s but failed to grow and did not spread to other organizations. Gradually, later projects became more holistically designed. Several large-scale regional projects in a variety of industries were started in the 1980s and continue today. They are supported by employers and unions and oriented toward participative dialogue among as many employees as possible. The aim is to achieve a participative–communicative infrastructure to facilitate a continuous process of adaptive changes and improvements in organizational effectiveness. This systemic Scandinavian approach has been compared to the Japanese total quality management, and both appear to have achieved a measure of success.

There are substantial opportunities for further research. For instance, one can hypothesize that holistically organized participative organizations would increase work motivation, reduce boredom, facilitate conflict resolution, and increase a sense of responsibility as well as compliance with safety measures. It is reasonable to predict that OP will continue to play a useful role in applied psychology and related social science areas.

5. CONCLUSION

Organizational participation has several names and operates in a variety of contexts. It can be formal (i.e., representative) or informal and ad hoc. It can be isolated from other managerial and HRM schemes, or it can be coupled with them in a systemic arrangement. Improvement of organizational performance is usually the objective. In the human resources model (Fig. 2), this is achieved because OP, where appropriate, uses experience and skills that in the absence of OP would remain dormant. Although the case for OP has been widely advocated in management programs and the literature, employment of OP has been moderate or low. There are several explanations for this contrast between espousal and implementation. It has been argued that OP is not independent of positive or negative environmental influence (Fig. 3), but research design rarely allows for this. Another view suggests that when OP is seen as empowerment of lower levels of organization, it can be assumed to be a threat to the "managerial prerogative." This interpretation is contested by those who argue that participation in the sense of contributing experience and skill does not abrogate the final responsibility of senior management to integrate the liberated competence in an effective decision-making process.

See Also the Following Articles

Competence at Work ■ Cooperation at Work ■ Decision Making ■ Industrial/Organizational Psychology, Overview ■ Organizational Culture and Climate ■ Organization Development ■ Organizational Socialization ■ Organizational Structure ■ Power, Authority, and Leadership

Further Reading

Cotton, J. (1993). *Employee involvement: Methods for improving performance and work attitudes*. London: Sage.

Heller, F. (2003). Participation and power: A critical assessment. *Applied Psychology: An International Journal*, *52*(1), 144–163.

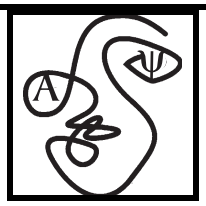
Heller, F., Pusic, E., Strauss, G., & Wilpert, B. (1998). *Organizational participation: Myth and reality*. Oxford, UK: Oxford University Press.

Ichniowski, C., Kochan, T., Levine, D., Olson, C., & Strauss, G. (1996). What works at work: Overview and assessment. *Industrial Relations*, *35*(3), 299–333.

Industrial Democracy in Europe Group. (1993). *Industrial Democracy in Europe revisited*. Oxford, UK: Oxford University Press.

Naschold, F., Cole, R., Gustavsen, B., & van Beinum, H. (1993). *Constructing the new industrial society* (Vol. 3). Assen/Maastricht, The Netherlands: Van Gorcum.

Pfeffer, J., & Veiga, J. (1999). Putting people first for organizational success. *Academy of Management Executive*, *13*(2), 37–48.



Organizational Socialization

Talya N. Bauer
Portland State University, Portland, Oregon, USA

1. Introduction
2. Stages of Socialization
3. Common Outcomes of Socialization
4. What Do Organizations Do to Facilitate Socialization?
5. How Do Insiders Aid in Socialization?
6. Newcomer Proaction
7. Socialization Research Methods
8. Conclusion
   Further Reading

GLOSSARY

accommodation The second stage of socialization, when newcomers learn how to fit into their new organization socially and task relatedly.

anticipatory socialization The first stage of socialization, when newcomers are thinking about joining the organization.

entry The point at which newcomers officially join the organization.

information seeking The process of new employees obtaining information, either overtly or covertly, about organizational life.

insiders Organizational members who are already established in the organization when a new employee joins.

longitudinal research Research that includes data collection at multiple points in time; this has typically occurred at 3, 6, 9, and/or 12 months postentry.

newcomers New employees who are undergoing the socialization process.

role management The third stage of socialization, when new employees become fully integrated into the organization.

socialization tactics The ways in which organizations train and orient new employees.

This article reviews what socialization is, why it matters, and thoughts on how to ensure its success. Relevant research findings are reported.

1. INTRODUCTION

Organizational socialization refers to the "learning" of what it is to be an organizational insider. It has been referred to as the process of "learning the ropes" of organizational life. By definition, organizational socialization deals with changes for newcomers. Unsuccessful socialization means that newcomers have not learned to fit into the organization either socially or with regard to the work of the job. Lack of socialization can result in high turnover rates and employee dissatisfaction. On the other hand, successful socialization can help organizations to retain valuable employees and may result in considerable cost savings. When socialization is effective, newcomers understand and adopt the organization's central values and norms. This guarantees better coordination among employees. Learning is key when looking at socialization. This article summarizes some of the important elements of organizational socialization and gives some practical implications for people in organizations.

©2004 Elsevier Inc.
All rights reserved.

2. STAGES OF SOCIALIZATION

Several scholars have proposed that socialization takes place in distinct stages. For example, the first stage is one of "anticipatory socialization." This stage includes all of the ways in which newcomers prepare themselves for entry into the organization as well as interaction they have with the organization such as recruitment, interviews, the selection process, and paperwork. This is followed by the "accommodation" stage. During this second stage, organizationally entry begins and newcomers tackle the challenges associated with learning to deal with the people as well as the actual work of doing the job. The final stage of socialization, which has been referred to in a variety of ways such as "role management," is when newcomers have mastered both the social and task-oriented portions of the job.

3. COMMON OUTCOMES OF SOCIALIZATION

Research on socialization has focused on newcomer adjustment and outcome variables. Adjustment variables include how much clarity newcomers have with their role, how accepted they feel, and how well they are adjusting to the organizational or career culture. Culture refers to the norms and operating guidelines within an organization, so these are not usually explicitly defined for new employees. Rather, they must discern them through interactions and questions. Other outcomes that have been studied include job attitudes, such as job satisfaction and organizational commitment, as well as performance and turnover. Research has shown that socialization is positively related to adjustment, job attitudes, job performance, and organizational culture and is negatively related to turnover.

4. WHAT DO ORGANIZATIONS DO TO FACILITATE SOCIALIZATION?

In Van Maanen and Schein's seminal article on socialization in 1979, these authors delineated six tactics that vary from individualist to institutional: the level of interaction with other newcomers (formal vs informal), the number of newcomers within a given cohort (collective vs individual), the order in which socialization takes place (sequential vs random), whether or not there is a specific time frame for socialization (fixed vs variable), how newcomers are trained (serial vs disjunctive), and whether or not newcomers are stripped of their old identity (investiture vs divestiture). All of these dimensions can be boiled down to a continuum with highly institutionalized approaches, where newcomers are socialized in a formal setting as a group and given clear information about the sequence and timing of the socialization process, and highly individualized approaches, where newcomers are socialized informally and given little explicit information about sequence or timing. Organizations vary in terms of these tactics. For example, in an extreme case, the military has very institutionalized formal tactics where individuals are brought in as large groups, the individuals go through formal training that is tied to specific goals and time lines, and individuality is not encouraged regarding how tasks are done.

5. HOW DO INSIDERS AID IN SOCIALIZATION?

Although a key focus on socialization research is newcomers, organizational insiders are important as well. Peers, supervisors, and mentors are often referred to as "agents" of socialization and are seen as playing an integral role in facilitating newcomer sense making. Insiders can facilitate socialization by providing newcomers with additional information, giving advice, clarifying job instructions, and giving social support.

Research has shown that the strength and nature of the newcomer–manager relationship relates to the success or failure of newcomer socialization. In general, research has also shown that job attitudes are positively related to the quality of newcomers' relationships with peers and the frequency of their communication with peers and that performance is related to social support from coworkers.

A significant discovery in organizational socialization research is how important mentors are for successful socialization. Mentors can provide many benefits such as teaching newcomers about the organization, offering advice, helping with job instructions, and providing social support. Mentors are also important because they can be a sounding board for questions that employees are too intimidated to ask their managers, and mentors can also provide the support structure for new employees to help them fit in both socially and politically. It is important for an organization to

recognize the importance of mentors when bringing new members into the organization. This can help with many things such as meeting and managing expectations better. Mentors can help new employees to learn about the organization more efficiently. Events and practices of the organization can be relayed with mentors simply through interaction.

6. NEWCOMER PROACTION

It has been argued that newcomers' personalities can influence the ways in which they see their roles and the challenges facing them. Information seeking also seems to affect organizational socialization. Studies have also demonstrated predictable patterns in the relative frequency with which newcomers seek various types of information. For example, newcomers tend to seek more technical information than social information, believe that there are costs associated with seeking information, and believe that seeking technical and referent information is the safest form of information seeking.

A key form of learning that newcomers need to gain is how to fit into the organization. Research has shown that a variety of factors influence how well newcomers learn to fit over time.

7. SOCIALIZATION RESEARCH METHODS

Socialization researchers tend to conduct longitudinal studies, use actual newcomers to organizations, and (often) use 3-month intervals (i.e., 3, 6, 9, and 12 months) to measure socialization-related variables, especially outcome measures. However, little empirical evidence has been offered to support the selection of these measurement periods beyond the fact that other studies have used them and found significant results using them. Often, the samples studied are college graduates or specific professionals such as nurses or accountants. Thus, what we know about socialization must be inferred from a relatively narrow set of occupations or job types studied.

8. CONCLUSION

The overall conclusions include the need for organizations to be proactive when socializing newcomers and the need for them to be aware of the surrounding conditions. Learning efficiently comes from training programs that may include mentoring, internships, and other ongoing training programs that are extremely important for socialization given that learning is key during the early stages. There is a need to understand how to effectively manage the changing workforce and how it affects the socialization process because the workforce is a key factor in determining employee commitment, performance, and retention and affects how people view the organization's culture. It is also necessary to have different socialization processes for different types of employees, such as permanent versus temporary (e.g., contingent, internships), and to examine where the particular organization is with regard to the changing workforce.

See Also the Following Articles

Employment Interviewing ■ Organizational Culture and Climate ■ Recruitment

Further Reading

Bauer, T., Morrison, E. W., & Callister, R. (1998). Organizational socialization: A review and directions for future research. In G. R. Ferris (Ed.), *Research in personnel and human resources management* (Vol. 16, pp. 149–214). Greenwich, CT: JAI.

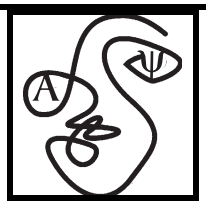
Feldman, D. C. (1976). A contingency theory of socialization. *Administrative Science Quarterly*, *21*, 433–452.

Fisher, C. D. (1986). Organizational socialization: An integrative review. In K. M. Rowland, & G. R. Ferris (Eds.), *Research in personnel and human resources management* (Vol. 4, pp. 101–145). Greenwich, CT: JAI.

Louis, M. R. (1980). Surprise and sensemaking: What newcomers experience in entering unfamiliar organizational settings. *Administrative Science Quarterly*, *25*, 226–251.

Morrison, E. W. (1993). Newcomer information seeking: Exploring types, modes, sources, and outcomes. *Academy of Management Journal*, *36*, 557–589.

Van Maanen, J., & Schein, E. H. (1979). Toward a theory of organizational socialization. In B. M. Staw (Ed.), *Research in organizational behavior* (Vol. 1, pp. 209–263). Greenwich, CT: JAI.



Organizational Structure

Lois E. Tetrick and Michael K. Camburn

George Mason University, Fairfax, Virginia, USA

1. Introduction
2. Traditional Dimensions
3. Modern Approach
4. Conclusions
   Further Reading

GLOSSARY

alternative employment relationships The use of different employment contracting to vary the number of hours worked, the length of employment relationship, and the combination of the two rather than the "traditional" full-time, long-term employee–employer relationship.

centralization The extent to which power is concentrated in the upper levels of management.

formalization The extent to which rules, regulations, and policies are written and enforced.

globalization The extent to which organizations cross national boundaries in producing and marketing their goods and services.

matrix structure An organizational structure characterized by the use of project teams in which employees may be members of multiple project teams; teams include expertise from multiple functional areas and exist only until the projects are completed.

organizational learning The process of acquiring new knowledge and skills that can be applied in novel ways.

organizational structure Mechanisms that determine the interactions among organizational members within the social system.

specialization The diversity of operations within an organization as well as the number of departments and positions within departments in an organization.

work teams Groups of employees that work to complete a specific task within the organization.

Organizational structure refers to the processes by which individuals within an organization interact with one another and with individuals outside the organization. The design of the organization, based on various organizing principles such as centralization, formalization, specialization, team structures, matrix structures, organizational learning, globalization, and alternative employment relationships, shapes individuals' behaviors and organizational functioning. These characteristics of the organization are interdependent, and contextual factors may moderate the effectiveness of organizational structure components in various environments both at the individual and organizational levels.

1. INTRODUCTION

The structure of an organization refers to mechanisms that control the interaction of organizational members in achieving organizational goals. The effects of organizing mechanisms on individual employees' behaviors and attitudes have received considerable attention, as have the effects of organizing mechanisms on

©2004 Elsevier Inc.
All rights reserved.

organizations' stability, growth, and adaptability. This article provides a summary of this research. First, it focuses on three traditional characteristics of organizational structure: formalization, centralization, and specialization. Then, it discusses modern approaches to organizing with an emphasis on teams, matrix organizing, organizational learning, globalization, and alternative employment relationships.

2. TRADITIONAL DIMENSIONS

In 1961, Burns and Stalker suggested that organizations were structured on a continuum ranging from mechanistic to organic. Mechanistic organizations reflected those in which power was centralized and control was achieved through rigid hierarchical communication channels and formal procedures. Mechanistic organizations are considered hierarchical, or tall, with multiple levels of management. Decision-making power is concentrated at the higher levels. Organic organizations, on the other hand, are more flexible and decentralized with open communication among organizational members across levels and minimal use of formal procedures. Organic organizations are heterarchical, or flat, with relatively few levels of management. Decision-making power generally is distributed across all levels in the organizations.

2.1. Centralization

Centralization refers to the concentration of power within an organization to the upper levels of the hierarchy, suggesting that the organization does not trust the employees to make decisions that are consistent with organizational goals. This concentration of power also has the potential to make employees more dependent on upper level managers for resources and information. Thus, in the centralized organization, control over resources and information is retained by the top management team and power resides in formal authority. Lower level personnel generally do not participate in decisions, even those regarding their own tasks.

2.1.1. Individual Consequences of Centralization

Employees working in more centralized organizations experience less autonomy as a result of the increased control over resources and information by individuals at the upper levels in the organization. According to Hackman and Oldham's Job Characteristics Model, reduced autonomy would be expected to result in lower motivational levels and ultimately in lower job performance. Also, the reduction in control resulting from lower levels of participation in the decision-making process can result in increased levels of stress. It would appear that decentralization is a more positive organizing mechanism from an individual perspective, although decentralization is not without some negative consequences. For example, in organizations where personnel decisions are decentralized, greater power is typically given to individuals' immediate supervisors in determining rewards and promotions. Because supervisors have greater power, there is an increased possibility that employees might feel as though they are being treated unfairly and might reduce their efforts. Depending on the quality of the relationship between an individual employee and his or her immediate supervisor, there might actually be a decline in job performance, especially discretionary aspects of job performance such as contextual performance and organizational citizenship behavior.

2.1.2. Organizational Consequences of Centralization

Centralization allows an organization, or at least the upper levels of the organization, to control lower level employees' behaviors so that they are consistent with organizational goals. Also, high degrees of centralization reduce the need for high levels of formal procedures because managers already directly monitor and advocate appropriate behaviors without needing written standards. However, as the organization increases its members and functions, the core group of top managers is less able to monitor all of the activities within the organization and more authority is shifted to departmental managers. Furthermore, centralization may constrain adoption of innovative ideas because lower levels in the organization have little decision-making power.

2.2. Formalization

Formalization refers to the extent to which there are formal rules, regulations, and policies in an organization. These formal rules and regulations control the behavior of employees, institutionalizing practices and procedures that have been successful in the past. If these procedures are followed, they provide consistency in decision making throughout the organization.

2.2.1. Individual Consequences of Formalization

The consistency of decision making provided by formal rules and procedures results in greater perceptions of organizational fairness and increased predictability within the work environment. This could result in increased job performance (especially contextual performance) and reduced job stress. In 2000, Schminke and colleagues suggested that the key to enhanced perceptions of fairness and predictability is not the extent of formalization but rather the extent to which the rules fit the organization's environment. For example, if the organization has a relatively stable and predictable environment, formalization matches the environment and is expected to have positive effects. However, if the environment is more fluid and dynamic, formalization might actually have negative consequences if the policies are perceived as unfair and/or too rigid. The gain in predictability from formalization can reduce job stress; however it can also lead to a sense of unnecessary red tape, a lack of control over the decision-making process, and a decreased sense of control, all of which can increase job stress.

2.3. Organizational Consequences of Formalization

Formalization results in standardization of processes and outputs, allowing an organization to maintain high levels of efficiency and productivity. For certain practices, such as selection and promotion, formalization helps the organization to avoid disparate treatment of applicants and candidates by standardizing procedures across all departments. This standardization of practices reduces the likelihood of legal actions against the organization. However, as Sitkin and Bies noted in 1993, promotion of a legalistic environment in an organization may actually undermine the social trust between managers and employees. Strict adherence to formal policies and procedures can limit the organization's adaptability and innovation much as in the case of high levels of centralization.

2.4. Specialization

Specialization refers to the diversity of operations performed within an organization. The organization may group similar areas of specialization into functional groups such as production, sales, and finance. Division of labor refers to the number of different job titles in the organization. In general, the greater the number of departments within the organization and the more positions within each department, the higher the degree of specialization. A high degree of specialization creates a more complex organization with multiple types of expertise.

2.5. Individual Consequences of Specialization

Higher levels of specialization result in employees perceiving a greater level of autonomy in an organization. This autonomy reflects the increased power of possessing unique skills and knowledge. Performing work that is not duplicated by others in the organization also enhances individuals' control over their own work.

2.5.1. Organizational Consequences of Specialization

Greater specialization within an organization indicates more divergence among managerial perspectives and interests. This divergence can lead to greater innovation, flexibility, and adaptability if there is not too high a level of formalization and centralization. However, upper level managers will often attempt to enforce similar standards across specialized subunits. This generally results in an increase, rather than a decrease, in formalization.

3. MODERN APPROACH

Modern organizations are becoming more decentralized and specialized. There appears to have been a transition to what Moch called more pluralistic political systems rather than hierarchically coordinated social systems. Organizations have become more diversified and heterarchical. They are concerned with the process of organizing and directing personnel. This organizing approach is characterized by growth, flexibility, and adaptability. Organized power hierarchies have weakened, resulting in greater decentralization and increased specialization. There also has been a shift from control through formalization to greater self-management, with modern organizations using a combination of formalization and flexibility to obtain reliable results while enhancing opportunities for innovation. Organizational structures are still characterized by the traditional dimensions, but a shift toward organizing has resulted in the increasing significance of a variety of new structural components. This section discusses several structural dimensions that

have emerged with the modernization of organizations: work teams, matrix structure, organizational learning, globalization, and alternative employment relationships.

3.1. Work Teams

One of the major factors associated with organizations shifting from the traditional hierarchical organization to a more diversified heterarchical organizational structure is the increased use of work teams. The increasing complexity of the business environment makes the use of work teams increasingly appropriate where diverse skills can be applied to strategic decision making. The use of self-directed teams makes fewer hierarchical levels necessary.

3.1.1. Individual Consequences of Work Teams

In the traditional approach to organizations, larger organizations increase the likelihood that employees will feel less valued. Use of work teams can functionally reduce the sense of largeness by embedding individual employees within smaller organizational units or teams. It also is likely that individuals will have a greater sense of accomplishment in completion of these team tasks rather than the smaller tasks that individuals might have in a non-team environment. However, if a work team has more members than is necessary, members might reduce their individual efforts. If the team is understaffed, on the other hand, members might experience low group potency, reduced motivation, and greater job stress.

3.1.2. Organizational Consequences of Work Teams

Organizing based on a work team structure allows organizations to undertake more complex projects. However, it is more difficult for upper level management to monitor and control performance in a team-based structure. Therefore, a specialized manager for each team is more effective and economical than a hierarchy of executives. The use of work teams also requires an organizational culture that trusts teams to work toward organizational goals relatively independently. This suggests that implementation of a team-based structure requires investment in training and development of teams and individual team members.

3.2. Matrix Structure

A matrix structure is a complex organizational structure that is based on the use of project-driven teams, generally requiring individuals with diverse functional expertise to be on the same teams. Project teams include multiple specialists, and each specialist may be part of multiple project teams. Members need to be highly skilled within their own discipline with strategic decision-making abilities and an ability to collaborate with others. Work teams, as described earlier, may work together indefinitely, but project teams in a matrix organization typically disband when the projects are completed. Matrix structures are intended to result in innovation, high task performance, and a collaborative shaping of organizational goals. Some automobile companies use a matrix structure during the development of new automobiles. Experts from a variety of functions, such as engineering, accounting, production, and marketing, all play a role in the success of new automobiles. Issues such as development, production, and sales are discussed, debated, and decided on by the team of experts, with each member bringing a unique and important perspective to the team.

3.2.1. Individual Consequences of Matrix Organization

The matrix organization involves considerable fluidity among organizational members, their assignments, and their social networks. Project team membership provides high levels of the core job characteristics. Thus, individuals would be expected to experience increased motivation and enhanced job performance. Furthermore, the social interaction among project team members has the potential to increase employees' social capital. However, working on multiple project teams with multiple managers can lead to role conflict and role ambiguity, both of which can result in job stress. Also, the relatively short duration of team membership and multiple team membership may reduce team members' identification with and commitment to each team.

3.2.2. Organizational Consequences of Matrix Organization

Project teams unite members with diverse skills and backgrounds that are expected to increase the quality of the projects and allow for innovative solutions if all members contribute. To some extent, they are the antithesis of stability in that they result in considerable movement within the organization by project team members. Matrix structures promote coordination across disciplines and functional areas within the organization, allowing a wide range of participation in

decision making and sharing of knowledge. However, communication across project teams is often not supported by the organization, and there may be power struggles among managers from different functional departments.

3.3. Organizational Learning

Adaptability requires learning, and as organizations are faced with competitive pressures, it is important that individual and organizational learning occur at a continual and rapid pace. Organizational learning is an active process in which knowledge is shared among organizational members, with information being shared across organizational subunits.

3.3.1. Individual Consequences of Organizational Learning

A structure that promotes individual learning promotes individual growth. If individuals are given opportunities to learn through training, they are likely to believe that the organization is supportive and committed to them as individuals and will reciprocate with increased organizational commitment and contextual performance. Also, when the organization provides an environment in which learning is valued and mistakes are tolerated, individuals will experience an increased sense of control over their work environment, thereby decreasing job stress and increasing well-being.

3.3.2. Organizational Consequences of Organizational Learning

The intended consequences of organizational learning are for the organization to be more innovative and adaptive to its environment. By acquiring and using new knowledge and skills, the organization is able to develop novel and more effective strategies for competing within the business environment. The effectiveness of organizational learning interventions varies across groups within an organization, depending on the interpersonal climate.

3.4. Globalization

Organizations have extended their domain from within a single country or region of the world to encompass multiple nations to reach a global market. There is no single standardized structure for multinational corporations. Their structures seem to reflect existing competitive environmental pressures and may vary across locations. Strategic decision making may be tightly centralized around senior management in the home country; it may be decentralized to the national level, allowing for strategic planning to be tailored for each national market; or there may be more globally dispersed resources, although strategic decision making is still likely to be centralized with formal committees to facilitate communication and coordination among subunits. In 1989, Bartlett and Ghoshal categorized these different structures of multinational corporations as multinational, international, global, and transnational based on the degree of interdependence. The multinational model is decentralized and nationally self-sufficient, the international model has sources of core competencies centralized and all others decentralized, the global model is centralized and globally scaled, and the transnational model is dispersed, being the most interdependent and specialized model.

3.4.1. Individual Consequences of Globalization

As organizations become more global, individual employees are expected to travel more, including potential long-term assignments in countries other than their countries of origin. Adjustment to new cultures can be a powerful learning experience and a source of stress. This can affect individual employees as well as their families, and organizations frequently provide acculturation experiences for employees and their families before overseas assignments. It also has been shown that maintenance of social networks within organizations is important for expatriates' adjustment to the assignments and for the expatriates' successful return at the end of the assignments. Another aspect of globalization is communication across national boundaries on a regular basis. For example, coordination of telephone conferencing may involve individuals being available at hours outside of their regular working hours. This can create a sense of organizational injustice if it is not handled appropriately.

3.4.2. Organizational Consequences of Globalization

Multinational organizations vary in the degrees of complexity of their structures. The basis for interdependence in these organizations results from the flow of products, technologies, and people among subunits and between subunits and headquarters. The degree of interdependence can range from low in multinational models, to medium in international and global models, to high in

transnational models. Formalization may be able to handle the necessary coordination when there are relatively low levels of interdependence; however, as the level and complexity of interdependence increase, formalization generally is insufficient. There is a need for informal means of coordination and control.

3.5. Alternative Employment Relationships

Much of the literature on organizational behavior has focused on full-time employees. New employment relationships have emerged to include core part-time employees, peripheral part-time employees, independent contractors, guest workers, and temporary workers. These various work arrangements reflect alternative processes in organizing personnel.

3.5.1. Individual Consequences of Alternative Employment Relationships

The advantage of the increasing number of employment relationships is that they give individuals greater choice in the number of hours they work, in the length of commitments to particular employers, and in opportunities for developing skills across work environments. In general, these effects are considered to have a positive effect on individuals' well being. A caveat to this positive aspect must be mentioned. If individuals are not able to match their employment arrangements with their desired employment arrangements, they may experience underuse, underemployment, and other negative consequences to their well-being. The existence of multiple work arrangements within an organization also can lead to status differences among employees, even though employees are essentially doing the same work. This can lead to perceptions of unfairness and discrimination.

3.5.2. Organizational Consequences of Alternative Employment Relationships

Alternative employment relationships provide organizations with the flexibility of engaging the services of individuals as needed. Maintenance of core employees provides stability to organizations, whereas peripheral employees provide organizations with the ability to expand and contract the size of their labor forces based on production demands. Also, part-time, temporary, and short-term contracting allow organizations to expand their labor pools to include individuals who might not be interested in full-time, long-term work arrangements. Thus, organizations can offer work arrangements that are consistent with their overall strategic decision making.

4. CONCLUSIONS

Organizational structures control social interactions within an organization. Centralization, formalization, and specialization have been shown to have effects on individuals' attitudes and behaviors as well as on organizations' stability, growth, and adaptability. Modern approaches to structure, such as the use of work teams, the adoption of a matrix structure, organizational learning, globalization, and the use of alternative employment relationships, reflect more of an organizing or process orientation to organizational structure, although they are not independent of centralization, formalization, and specialization. There are positive and negative aspects of all of these aspects of organizational structure, and context is important in understanding the choice of various aspects of organizational structure. It is recognized that structure should follow strategy so that the impact of the structure is consistent with the organization's strategy. However, as Burns and Wholey reported in 1993, adoption of a particular organizational structure may be explained somewhat by organizational characteristics, but it appears that what other organizations are doing in the local network influences the choice of a particular structure.

See Also the Following Articles

Industrial/Organizational Psychology across Cultures ■ Industrial/Organizational Psychology, Overview ■ Organizational Culture and Climate ■ Organizational Diagnosis ■ Organizational Participation ■ Organization Development

Further Reading

Bartlett, C., & Ghoshal, S. (1989). *Managing across borders: The transnational solution*. Boston: Harvard Business School Press.

Burns, T., & Stalker, G. M. (1961). *The management of innovation*. London: Tavistock.

Burns, L. R., & Wholey, D. R. (1993). Adoption and abandonment of matrix management programs: Effects of organizational characteristics and interorganizational networks. *Academy of Management Journal*, *36*, 106–138.

Coyle-Shapiro, J. A-M., Shore, L. M., Taylor, M. S., & Tetrick, L. E. (2004). *The employment relationship: Examining psychological and contextual perspectives*. Oxford, UK: Oxford University Press.

Easterby-Smith, M., & Lyles, M. (Eds.). (2003). *The Blackwell handbook of organizational learning and knowledge management*. Malden, MA: Blackwell.

Hackman, J. R., & Oldham, G. R. (1980). *Work redesign*. Reading, MA: Addison–Wesley.

Kostova, T., & Roth, K. (2003). Social capital in multinational corporations and a micro–macro model of its formation. *Academy of Management Review*, *28*, 297–317.

Mintzberg, H. (1993). *Structure in fives: Designing effective organizations*. Englewood Cliffs, NJ: Prentice Hall.

Moch, M. (1976). Structure and organizational resource allocation. *Administrative Science Quarterly*, *21*, 661–674.

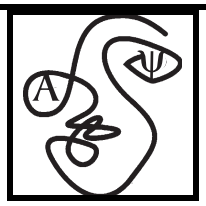
Pearce, J. A., II, & David, F. R. (1983). A social network approach to organizational design–performance. *Academy of Management Review*, *8*, 436–444.

Rousseau, D. M. (1997). Organizational behavior in the new organizational era. *Annual Review of Psychology*, *48*, 515–546.

Schminke, M., Ambrose, M., & Cropanzano, R. S. (2000). The effect of organizational structure on perceptions of procedural fairness. *Journal of Applied Psychology*, *85*, 294–304.

Sitkin, S. B., & Bies, R. J. (1993). The legalistic organization: Definitions, dimensions, and dilemmas. *Organization Science*, *4*, 351–354.

Tetrick, L. E. (1997). Macro-organization: Organization structure. In J. M. Stellman (Ed.), *ILO encyclopaedia of occupational health and safety* (4th ed., pp. 34.35–34.36). Geneva, Switzerland: International Labor Office.



Organization Development

W. Warner Burke

Teachers College, Columbia University, New York, New York, USA

1. An Organization Development Case
2. Primary Characteristics of OD
3. Roots of OD
4. Related Theories That Underlie OD
5. OD as a Process of Change
6. Understanding Organizations: The Diagnostic Process
7. Four Models Commonly Used by OD Practitioners
8. Planning and Managing the Consequences of Change
9. Does OD Work?
10. OD Values and Considerations for the Future of the Field

Further Reading

GLOSSARY

action research Collecting relevant information (data) about a given issue or problem (the research aspect) for the express purpose of providing ways to help resolve the issue or solve the problem (the action aspect).

appreciative inquiry Acts of exploration and discovery (asking many questions) for the purpose of recognizing the best in people (e.g., affirming strengths, potential, and successes).

cognitive map A thought process, mental exercise of establishing and tracing paths throughout an organization to help clarify and take corrective actions for organizational improvement and change.

development Growth, advancement, strengthen, and/or to bring an idea into action.

group unconscious A theory emanating from Freudian psychoanalysis that argues that a group of people has a collective unconscious analogous to an individual's unconscious.

information technology system The use of computers for transporting information throughout an organization; system, in this case, means that all organizational members are (or can be) connected electronically and therefore form a communicative whole.

interdependent multivariate systems Statistical term that is used to convey the notion that all organizational variables and components are interconnected and form a whole; the statistical method is a way of numerically demonstrating this idea.

intervention The act of intervening, coming between two or more parties in an organization—boss and subordinates, two or more departments, divisions, or business function/units, an individual and the larger group of which he or she is a part, or the organization and another organization (e.g., acquisition or merger)—for the purpose of initiating or furthering organizational change and development.

meta-analysis A statistical procedure for comparing the results of many studies of like content.

multirater feedback Based on specific behavioral ratings, a process of providing information about a person's leadership, management of others, or simply a range of individual behaviors, whereby a comparison is made of how one rates oneself with how others rate him or her on the same set of behaviors.

nonlinear complex systems theory Originating from chaos theory in the physical sciences; the proposition is that organizational behavior is affected by many variables that interact powerfully and interdependently but not in a single direction; cause and effect therefore is rarely as simple as it sounds (i.e., being a "one-way street").

norm A standard or rule of conduct to which people conform (e.g., a dress code).

©2004 Elsevier Inc.
All rights reserved.

open system framework A scheme or model usually in a two-dimensional arrangement that depicts an organization as having input (external environment), throughput (the operation of the organization), and output (performance) with a feedback loop connecting performance with the external environment; any organization is dependent on its external environment for survival, as in a living organism, yet through performance has an impact on and influences its external environment.

organization A sociotechnical system that has some purpose in society; sometimes it is only a social system such as a club.

refreezing Once organization change has occurred, a process of stabilizing and institutionalizing the new behavior that supports the change.

self-directed groups Organizational units that conduct their work without a leader or manager who has formal authority; group leadership occurs but usually in an informal and rotating fashion.

sensitivity training Based on interacting with and learning from others in a small group, a process of helping individuals to develop greater self-awareness and become more sensitive to how they affect and are affected by others.

sociotechnical system The principle that a work organization is both a technical system—the content or nature of the work itself—and a social system—the people involved and how they interact to get the work done; to overlook either system is to misunderstand the organization.

transactional That aspect of an organization involving daily operations, including transactions between organizational members and between them and those the organization serves.

transformational That aspect of an organization involving the larger system in the context of long time frames; its purpose, mission, strategy, culture, and senior leadership and how these organizational variables interact with the external environment and with one another.

unfreezing The act of changing ("melting away") current employee and managerial behavior that no longer leads to optimum organizational efficiency and effectiveness, with the act being education and training, the introduction of new technology requiring different work habits, providing evidence from the external environment that a new strategy with different products and services is required to ensure long-term survival, or all of these.

Organization development (OD) is about changing an organization—an institution, a company, a government agency, a community service agency, a university, a school, or even a church, synagogue, or mosque. Organization therefore is broadly defined—that is, a sociotechnical system that has some purpose in society. Sometimes, however, it is only a social system such as a club. The term development means growth, advancement, strengthen, or to bring an idea into action. By implication at least, development is about change. The change may be gradual or sudden but change nevertheless. Organization change can happen in a number of ways—by an election or appointment of a new leader; coercion via some civil disturbance; acquisition by a larger, more powerful organization; introduction of a new technology that immediately replaces the old; etc. For OD, organization change occurs by way of the application of behavioral science—that is, using knowledge from psychology, sociology, and cultural anthropology to develop and hence change the sociotechnical system.

1. AN ORGANIZATION DEVELOPMENT CASE

The organization development (OD) example described here began with the merger of two banks in a large metropolitan area of the United States. The two organizations were very similar with respect to number of employees, length of existence (both founded in the mid-19th century), and annual revenues, but they were highly dissimilar regarding their respective cultures. For example, one focused much more on cost-effectiveness as the way to a profitable bottom line, whereas the other emphasized revenue growth via developing new markets and services. The latter's motto seemed to be "you have to spend money to make money." In other words, much work would be needed to successfully merge the two banks.

The CEO of the newly merged bank sought consultation—one for strategic planning and another for OD. The two consulting firms that were brought in worked in tandem but focused on different initiatives to facilitate the merger and to foster change. The OD firm's consultation began with three primary initiatives. The first initiative was to conduct interviews with the top 19 executives of the merged bank. These interviews consisted of broad, open-ended questions, such as "What is going well regarding the merger?" and "What is not?" The interviews served as a data-gathering source for the second initiative, which was an organizationwide survey. The survey was conducted 6 months into the merger to assess the general state of morale and, specifically, to determine the degree of understanding and opinions about the bank's strategy, its leadership, the newly merged culture,

organizational structure and systems, and the adequacy of management in general. The third initiative was to discard the two previous banks' mission statements and craft a new one for the merged bank. To accomplish this initiative, a task force of 16 employees was formed that represented a cross section of the bank. Working in pairs, the task force sought input from the various parts of the bank they represented, eventually involving several hundred employees. They used the input from the expanded group of employees to craft the initial draft of a company mission statement. After a number of revisions the task force presented its statement to senior management and later to the board of directors. A final version was unanimously agreed on by both senior management and the board. In fact, the task force received a standing ovation for their work from the board of directors.

Results from the organizationwide survey using a 5-point Likert scale showed that most employees were positive about the merger (example of survey question: "To what extent have mergers and acquisitions made the bank more efficient?") and their leadership, at least at the very top; clear about the new strategy (e.g., "To what extent are the bank's products and services competitive with the top tier of banks in our geographical region?"); and strongly supported the new mission statement (e.g., "To what extent does the bank's mission include valuing cultural differences?"). On the other hand, employees were somewhat confused about culture (e.g., "To what extent is it safe to say what you feel?") and structure, especially with respect to accountability; critical of the reward system, especially compensation and incentives (e.g., "To what extent does the bank have rewards and benefits necessary to attract and retain the very best people?"), information technology (e.g., "To what extent are the computer systems that you use advanced and efficient?"), and the budgeting process; and quite negative about leadership below the very top five or so executives and highly critical of management (e.g., "To what extent does your supervisor provide people with the information they need to do their jobs effectively?") overall.

The survey results led to further initiatives and interventions. The information technology system was strengthened, the reward system improved, and a major initiative was launched to address the deficiencies in leadership and management beginning with the top 125 executives and managers. These programs for leadership and management emphasized self-awareness for individual development via multirater feedback, personality and leadership assessment, and coaching.

Approximately 1 year after the first organizational survey, a second survey was conducted. Responses to 125 questions resulted in two questionnaire items remaining the same as those from the first survey and all remaining items showed higher ratings. Although it cannot be scientifically proven that the desired changes for the bank did indeed occur, the bank's executives, managers, and employees believed that positive change had occurred. The results of the survey were interpreted within the framework of an overall organizational model of performance and change. The model included dimensions such as external environment, mission and strategy, culture, leadership, structure, systems, climate, motivation, and performance.

2. PRIMARY CHARACTERISTICS OF OD

The previous case helps to illustrate what constitutes an OD approach to organization change. The following characteristics, when applied to the case, help exemplify the essence of OD.

2.1. Relevance

OD must be relevant to the organization's purpose and to what its employees actually do every working day. In the bank case, the OD practitioner worked initially on the organization's mission and ensured that this work linked directly to the bank's business strategy.

2.2. Sequential

OD practice involves a sequence of activities. After entry and contracting (reaching an agreement about the consulting services to be provided and when), the OD practitioner then gathers data about the organization and its leadership/management, determines a diagnosis about what the data mean, provides feedback to the client organization, and facilitates collaboratively with the client several interventions (initiatives) that help to address issues and needs for change embedded in the data and the diagnosis. After the interventions are launched, the next step in the sequence is to gather data again in order to evaluate progress. In the case,

1. Contracting occurred with the CEO and his immediate direct-report executives.
2. Data gathering took the form of executive interviews, which in part set the agenda for further data gathering via the initial organizationwide survey.
3. Based on the initial interviews of the top 19 executives, feedback was provided to them that led to off-site meetings facilitated by the OD consultant to address issues highlighted from a summary of the 19 interviews. Further larger system feedback was later provided for all managers in the bank and subsequently to all employees from the organizationwide survey.
4. Interventions from the initial feedback consisted of top executive team building, changes in the information technology and reward systems, and the launching of leadership/management development workshops.
5. Further data gathering and evaluation of the change initiatives (interventions) up to that point were conducted using a second organizationwide survey.

2.3. Systemic

OD is a system approach to organization change. In the bank case, this was manifested in the construction and feedback of the data from the survey. The 125 items comprising the survey were arranged in 12 categories—that is, according to the 12 boxes in the Burke–Litwin causal model of organization performance and change. This model is an open-system framework (i.e., input–throughput–output with a feedback loop connecting output to input) and consists of the following categories: external environment, mission/strategy, leadership, culture, structure, management practices, systems, work-unit climate, job–person match, individual needs and values, motivation, and performance. The survey feedback in the bank case was in the form of these 12 categories with priorities for action steps according to the arrangement of the boxes in the model in Fig. 1.

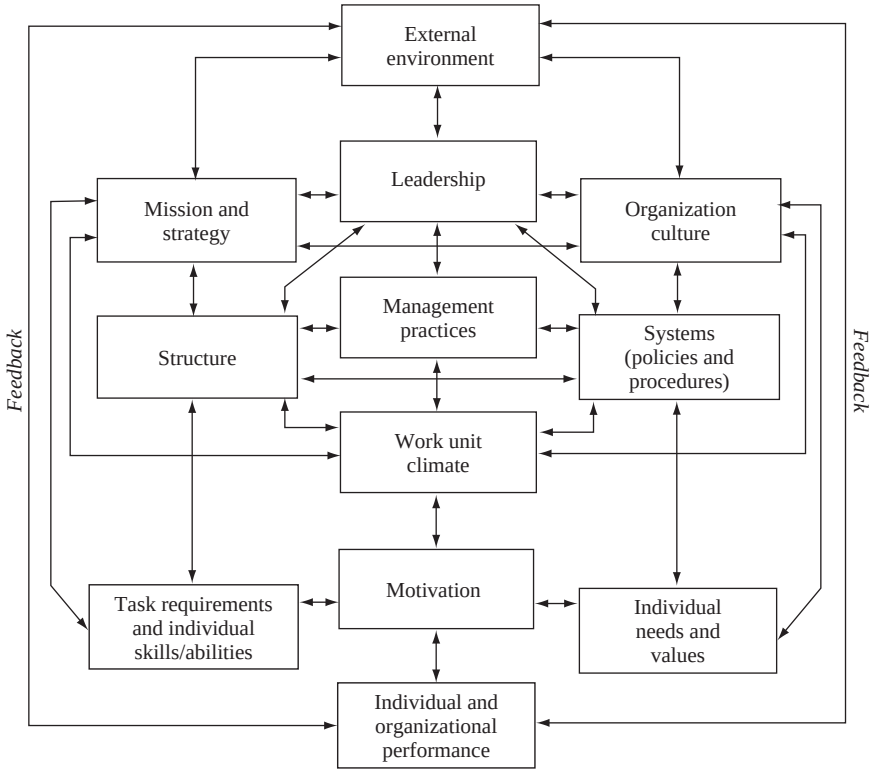


FIGURE 1 The Burke–Litwin model of organizational performance and change. From the Burke–Litwin model of organizational performance and change. © 1987, 1992 by W. Warner Burke, Associates, Inc.

2.4. Research Based

OD is steeped in the tradition of action research. This means that whatever action for change is taken, it is based on data, usually both qualitative and quantitative. OD interventions are "organic"—that is, based on what the data and the diagnosis reveal. Therefore, the action is based on the research (research loosely defined).

2.5. Humanistic

OD practice typically conforms to a set of values grounded in humanism and often takes the form of

- Involvement of organizational members in decisions that directly affect them
- Respect for differences among people
- Openness and honesty in communication
- A strong belief that the solution to organizational problems often lies within the knowledge and experience of people in the organization who have lived with the problem daily.

Usually this knowledge is either unspoken, perhaps due to an implicit norm not to disagree with the boss or with one another or the problem is a "sacred cow," or the knowledge is tacit—that is, what may be known is not forbidden to discuss but, rather, is implicit and difficult to articulate. The OD practitioner's role and function in these instances is to bring to the surface what organizational members are fearful of discussing or to help make explicit what is tacit, implicit. The previous four values are meant to be representative, not exhaustive.

3. ROOTS OF OD

Where did OD come from? Three primary streams of change methodologies and three institutions provide the basis for the emergence of OD.

3.1. Sensitivity Training

Evolving from the work of Kurt Lewin and colleagues during the 1940s, this form of training concerned the application of knowledge regarding group dynamics and interpersonal relations for individual learning. The learning occurred via feedback from others as to how one's behavior affected these other group members. Learning also came from an analysis of interpersonal relations in small groups in which the primary source of information or data from which to learn was the behavior and feelings of the group members. In the late 1950s and early 1960s, sensitivity training was employed in organizations to enhance interpersonal, managerial, and team effectiveness. These activities became the forerunner of what is a cornerstone of OD today—team building.

3.2. Sociotechnical Systems

Most any organization is at the same time a technical system and a social system, and to ignore either one during a change effort is to invite failure. Trist and Bamforth, both professional staff members of the Tavistock Institute, served as consultants to a coal mining company in the northern part of England. The company had made a technological change regarding the way coal was mined and instead of the change increasing productivity as management expected, morale of the workers plummeted and productivity declined. After studying the situation, Trist and Bamforth noted that the technological change had changed the social system. The miners could no longer work in small groups as before but were forced to work alone, resulting in a greater division of labor. The consultants explained to management what the real problem was and how to make another change that would not result in a return to the old method but modify the work process so that it would be more composite (i.e., less divided) and help to retain the social patterns that were so important to the miners. This way of thinking and changing organizations led to a number of interventions that are used by OD practitioners today, such as job redesign, self-directed work groups, and organizational redesign, especially in manufacturing companies.

3.3. Survey Feedback

Although not widespread, using a questionnaire to assess worker morale, attitudes, and perceptions in the 1950s was not new. What often occurred with the use of an organizational survey, however, was that management would retain the data and rarely report the results to those who responded to the questionnaire in the first place. Psychologists, particularly Rensis Likert and Floyd Mann at the University of Michigan's Institute for Social Research, found that surveys did not work very well because of this lack of "reporting back." Sometimes as a result of conducting a survey matters got worse. In other words, when asked to answer a questionnaire, employees thought that management

cared about what was on their minds. However, when nothing happened as a result of the survey, employees would become dissatisfied, even cynical, and productivity and morale would decline to levels lower than those before the survey was conducted. These Michigan psychologists believed that reporting the results to the people who answered the survey in the first place was a necessary step toward making the process a more positive experience. Necessary, they believed, but not sufficient for bringing about positive organization change. Mann, in particular, developed a way of stimulating action in an organization as a consequence of the survey. His procedure was to report the survey results to intact work and management teams—that is, to people who worked together. These work units would then analyze the results with respect to their own domain of work and responsibilities and then plan and subsequently take steps toward solving the problems identified in the survey. This procedure may not be groundbreaking today, but in the 1950s it was rare indeed. The procedure was called survey feedback simply because the results were "fed" back to managers, staff, and work units in a structured manner and the process set into motion action for resolving problems at the work unit level (i.e., where the knowledge resided about what to do). Survey feedback is common practice in OD today.

Although these three primary historical roots of OD differ, they all have in common the action research methodology and mode of thinking regarding organization change. It should also be recognized that theory, research, and methodology from industrial/organizational psychology have contributed significantly to the practice of OD.

4. RELATED THEORIES THAT UNDERLIE OD

There is no all-encompassing theory of OD. There are a number of "mini" theories that help to undergird the practice, however. The following sections provide a simple listing of these theories. They are grouped according to organizational level and emphasis of OD practice.

4.1. Individual Level

4.1.1. Need Theory of Motivation

Maslow's theory concerns a hierarchy of human needs that begin with physiological or creature comfort needs (air, water, safety, etc.), include security, belongingness, and ego status, and conclude at the top of the hierarchy with self-actualization. Herzberg's theory is more about job satisfaction and dissatisfaction than about needs per se. Factors that reduce dissatisfaction he called "hygiene" or preventive, such as fringe benefits, good working conditions, and salary. Factors enhancing job satisfaction he called motivators and include recognition, challenging work, and autonomy. OD practice that relies on these theories includes job enrichment, quality-of-work-life projects, empowerment interventions, and career development.

4.1.2. Expectancy Theory of Motivation

This theory is cognitive in nature and consists of individuals' expectations that (i) their efforts will lead to results, (ii) if enhanced these efforts will lead to rewards, and (iii) these rewards will be worth it to them. OD practice associated with this theory includes work redesign and modifying performance appraisal and/or systems.

4.1.3. Job Satisfaction

Related to Herzberg's theory, this theory posits that there are three primary psychological states that affect job satisfaction: experienced meaningfulness of the work, experienced responsibility for the work and its outcomes, and knowledge of results (i.e., performance feedback). OD practice that relies on this theory includes primarily job redesign, job enrichment, and performance feedback procedures.

4.1.4. Positive Reinforcement

Although Skinner's research and theory were more about individual animal and human behavior in the broad context of learning and change and not about organizational behavior per se, his principles can be applied to organizational reward systems. The primary emphasis is therefore on incentives and the belief that partial and variable reinforcement of behavior is more lasting and effective than continuous reinforcement. OD practice that relies on this theory includes modifying a reward system in general and putting into place some form of an incentive system.

4.2. Group Level

4.2.1. The Group as the Focus of Change

Applying concepts from the science of physics, Lewin analyzed human behavior in terms of "forces" impinging on an individual. These forces might be group norms to which the individual is expected to conform, supervisory pressure to exert more effort on the job, or deeply held beliefs and values emanating from the larger culture about what is the right thing to do. Changing norms and values is therefore tantamount to changing an organization's culture. OD practice that relies on this theory includes team building, force field analysis, decision-making processes, and ultimately culture change (i.e., changing conformity patterns).

4.2.2. Changing Values through the Group

The work of Argyris is broad and thus difficult to reduce to one theory. Appropriate for OD are primarily ideas about change, particularly the motivation to change. The premise is that when individuals (think managers) are confronted with differences between what they espouse and what they actually do, they will be motivated to reduce the gap and as a result change their behavior. OD practice that relies on this theory includes "gap analysis"—that is, surfacing the incongruence between what managers say, the words they use (Argyris calls it espoused theory), and what their actions actually become (theory in action), especially decision making—and team building. Also, in collaboration with Schön, Argyris helped spawn another OD activity that is more at the larger system level—organizational learning.

4.2.3. The Group Unconscious

Steeped in psychoanalytic theory, Bion took Freud's thinking one step further and argued that a group has a collective unconscious much like an individual's. This collective unconscious revolves around the group leader, whether formal or informal, and the group's motivation, albeit below conscious awareness, is to replace the leader irrespective of his or her performance. OD practice that relies on this theory includes group behavior diagnosis and team building, self-directed work groups, and sociotechnical systems interventions.

4.3. Larger System Level

4.3.1. Participative Management—One Best Way

A normative theory, Likert's position was that any large organization could be considered as one of four possible systems. System 1 was authority centered, top-down, and unilateral; system 2 was a benevolent autocracy; system 3 was consultative (i.e., ask subordinates for their ideas and opinions but with the manager reserving the right to make the final decision); and system 4 was participative management with an emphasis on consensual decision making. Likert's argument was that system 4 was the best way to manage an organization. OD practice that relies on this theory includes survey feedback and team consensus decision making (i.e., change to a participative form of management).

4.3.2. It all Depends

Opposite to a normative way of thinking, Lawrence and Lorsch stated that how an organization should be managed and designed depends on its relationship to its external environment (i.e., the organization should be as responsive as possible to external forces). Their key concept was interface—three of primary importance: the organization–external environment interface, intergroup interfaces within the organization, and the implicit relationship between the organization and the employee (the psychological contract). OD practice that relies on this theory includes organizational structure redesign, intergroup conflict resolution, and cross-functional groups for purposes of integration.

4.3.3. The Organization as a Family

Levinson considered top management as replicating the nuclear family, with the CEO being the father and the remaining executives acting as siblings, with rivalry among them being the dominant theme. Not surprisingly, Levinson's work has been most applicable to family-owned businesses. OD practice that relies on this theory includes succession planning in a family-owned business, coaching the CEO, and team building for the top executive group.

Table I summarizes the minitheories discussed in the previous sections.

TABLE I
Summary of Primary OD Theorists According to Their Perspectives, Emphases, and Applications[a]

Perspective	*Theorist*	*Emphasis*	*Application*
Individual	Maslow and Herzberg	Individual needs	Career development, job enrichment
	Vroom and Lawler	Individual expectancies and values	Reward system design, performance appraisal
	Hackman and Oldham	Job satisfaction	Job and work design, job enrichment
	Skinner	Individual performance	Incentive systems, reward system design
Group	Lewin	Norms and values	Changing conformity patterns
	Argyris	Interpersonal competence and values	Training and education
	Bion	Group unconscious, psychoanalytic basis	Group behavior diagnosis
System	Likert	Management style and approach	Change to participative management
	Lawrence and Lorsch	Organizational structure	Change contingent on organizational environment
	Levinson	Organization as a family, psychoanalytic basis	Diagnosis of organization according to familial patterns

[a]*Source:* W. W. Burke, *Organization development: A process of learning and changing* (2nd ed.). Copyright 1994 Addison-Wesley Publishing Co. Reprinted from p. 53 by permission.

5. OD AS A PROCESS OF CHANGE

As stated earlier, OD is based on action research as an approach (Fig. 2) and immersed in the Lewinian tradition exemplified by the following phrase: No action without research and no research without action.

5.1. Levin's Three-Step Change Procedure

Lewin provided a simple (although not easy to execute) three-step procedure for bringing about change. The following are the three steps:

1. Unfreeze the current level of behavior
2. Action that changes the system from its original state to a new level of behavior
3. Refreeze, meaning to secure the new behavior against further change or "slippage"

5.2. Schein's Elaboration on Levin's Three-Step Procedure

In 1987, Schein developed a useful elaboration of the three steps. He refers to them as stages rather than steps to emphasize the point that they overlap (i.e., they are not discrete). Schein's expansion of the three stages is presented in the following sections.

5.2.1. Stage 1: Unfreezing

Schein defined this stage as creating motivation and readiness to change. The following are his three ways of unfreezing:

1. Disconfirmation of the way things are, which leads to dissatisfaction and a need for change (e.g., costs are too high, morale is too low, and productivity is declining).
2. Induction of guilt or anxiety by establishing a gap between the current state and what would be much better. In the face of this kind of gap, most people become motivated to reduce it.
3. Creation of psychological safety to help deal with disconfirmation, guilt, and anxiety so that people can take the required action. The action needs to be associated with feelings of accomplishment and worthiness instead of humiliation or a loss of self-esteem.

5.2.2. Stage 2: Changing

Changing is the movement toward a new state that Schein believes can be accomplished via two processes. One process is the identification of organizational

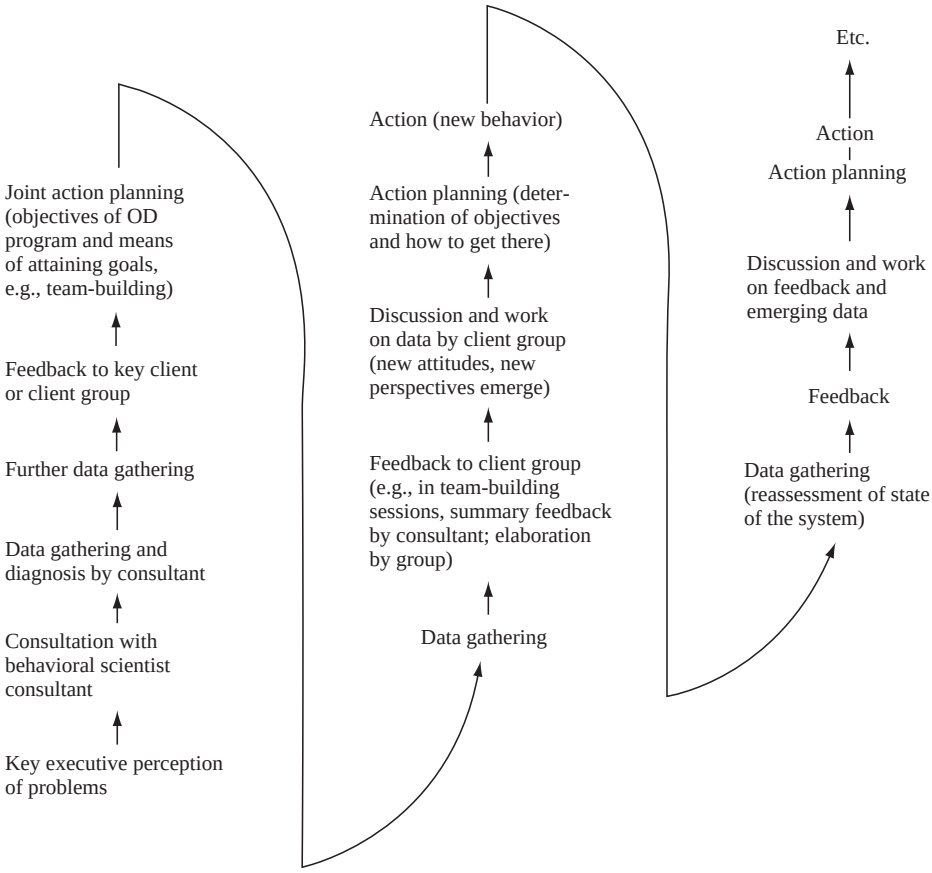

Key executive perception of problems
Consultation with behavioral scientist consultant
Data gathering and diagnosis by consultant
Further data gathering
Feedback to key client or client group
Joint action planning (objectives of OD program and means of attaining goals, e.g., team-building)
Data gathering
Feedback to client group (e.g., in team-building sessions, summary feedback by consultant; elaboration by group)
Discussion and work on data by client group (new attitudes, new perspectives emerge)
Action planning (determination of objectives and how to get there)
Action (new behavior)
Data gathering (reassessment of state of the system)
Feedback
Discussion and work on feedback and emerging data
Action planning
Action
Etc.

FIGURE 2 Action research model for organization development. From W. L. French, *Organization development: Objectives, assumptions, and strategies*, © 1969 by the Regents of the University of California. Reprinted from the California Management Review, 12(2), 26, by permission of the Regents.

members with a new role model, new leader, or perhaps the OD consultant to obtain a more attractive point of view and perspective. The second process is to scan the environment for new but relevant information that can facilitate the acquiring of a different perspective and point of view. This new information can come from a research project, outside expertise, or people in another organization who have gone through a similar organization change.

5.2.3. Stage 3: Refreezing

Schein conceptualizes this stage in two ways: the individual and relations with others. Refreezing with respect to the individual organizational member is a matter of feeling comfortable with the change—that it fits well with one's self-concept, values, and new modes of behavior. Relations with others concerns also achieving a level of comfort. As one's behavior changes, the new behavior affects others differently than in the past. Will this be okay? Will others also change and work toward congruence of these new behaviors with one another? Reaching and maintaining these new levels is what Schein (and Lewin) means by refreezing.

5.3. Lippitt *et al.*'s Five-Step Elaboration

Another elaboration of Lewin's three steps was developed in 1958 by Lippitt *et al.*, who expanded the three steps into five phases:

1. Development of a need for change (Lewin's unfreezing)
2. Establishment of a change relationship between the client and the consultant(s)

3. Working toward change (Lewin's action movement step)
4. Generalization and stabilization of change (refreezing)
5. Achieving a terminal relationship between the client and consultant(s)

Lippitt and colleagues viewed an organizational change effort from the perspective of the consultant, which they referred to as the "change agent." To summarize, four models or approaches to organization development have been briefly described: action research, Lewin's three steps, Schein's elaboration of Lewin's steps, and Lippitt *et al.*'s expansion. Figure 3 provides an overall summary and integration of these approaches.

Since these approaches overlap, they can be considered together in the form of an overall approach for OD that consists of the following elements:

- An outside consultant or change agent (the consultant may be an employee of the organization but would still be an "outsider" to the group in the organization being served)
- Data gathering from the client system (either the total system or some part that is relatively distinct—a

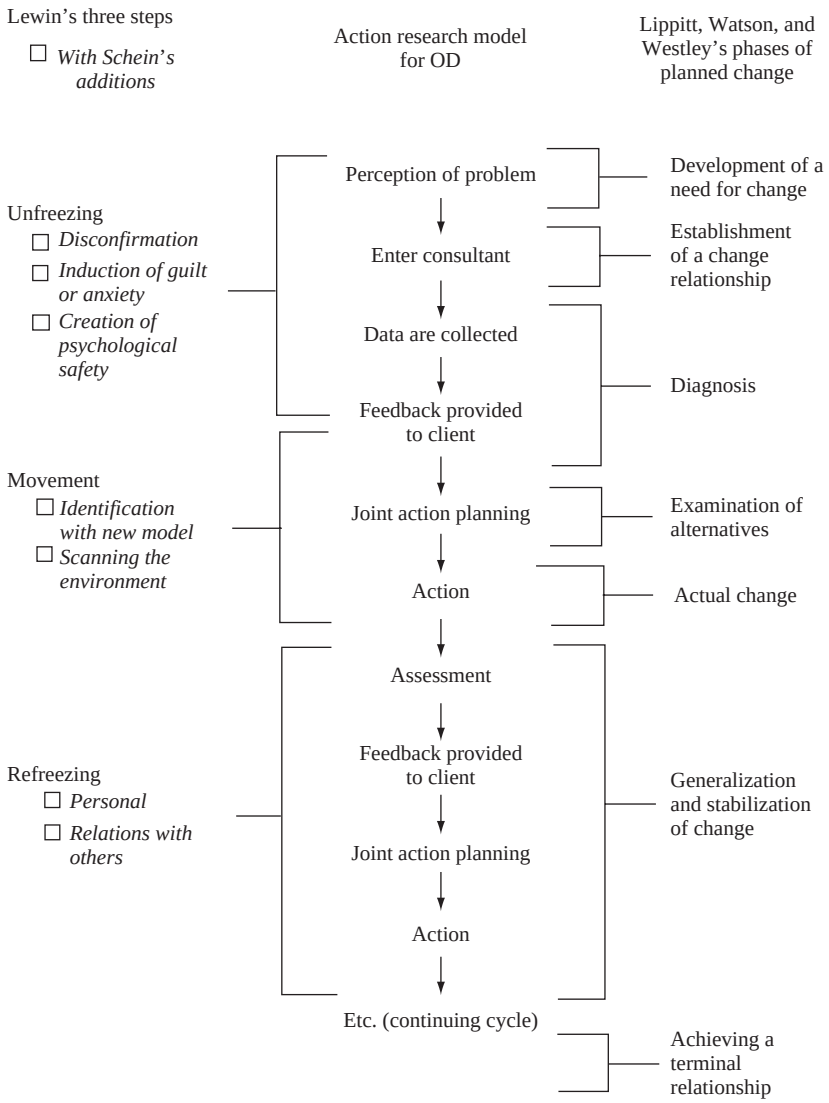


FIGURE 3 Comparison of the four models of change. From W. W. Burke, *Organization development: A process of learning and changing* (2nd ed.). © 1994 Addison-Wesley Publishing Co. Reprinted from p. 64 by permission.

division, a business unit, etc.) for the purpose of diagnosis (i.e., to understand in as much depth as possible the inherent nature of the system and its culture)

- Determine the major domains that need to be changed and problems to be solved
- Report this diagnosis to the client system so that appropriate action can be planned in a collaborative manner between the consultant (OD practitioner) and the client system
- Implementation of the planned change that is based on valid information in the form of interventions into the system that is usually conducted by the clients but with the continuing help of the consultant
- Institutionalization of the change

In short, the seven phases and processes of OD practice are as follows:

1. Entry: Establishing initial contact and a relationship with the client
2. Contracting: Agreement about the work to be done together (consultant and client); for example, who will do what and when
3. Diagnosis: Consists of data gathering and making sense of the information that has been collected (an organizational model helps with the latter process)
4. Feedback: Reporting to the client the summarized data and what the data mean
5. Planning change: Establishing action steps to accomplish the change
6. Intervention: Initiatives and actions that lead to change
7. Evaluation: Measures that help to determine if the change is taking effect

6. UNDERSTANDING ORGANIZATIONS: THE DIAGNOSTIC PROCESS

Methods for data gathering in the OD process consist of (i) structured interviews, (ii) questionnaires either to embellish interviews (or vice versa) or to collect data in the form of an organizationwide survey, (iii) archival information (e.g., statements of mission and strategy, annual reports, and planning documents), and (iv) observations by the OD practitioner (e.g., paying attention to artifacts in the organization, sitting in on key management meetings, and listening carefully to the language of organizational members). Often, these forms of data gathering produce more information than either the consultant or the client can deal with. Moreover, these data in their "raw form" are not obvious with respect to meaning. After all, the data are more about symptoms than causes of problems and issues. Thus, a framework or organizational model for summarizing and understanding a large amount of data can be useful. To be useful, the data need to be treated in organizational terms. OD represents a systems approach to change and a model can help to translate raw data into systems language. Thus, a model needs to be an effective representation of an organization.

An early organizational model was developed by Harold Leavitt in 1965. His diamond-shaped framework consisted of the following four concepts positioned at the corners of the diamond with interconnecting lines: structure, technology, people, and task. His model helped to set the stage for thinking about organizations as interdependent multivariate systems. In fact, the popular McKinsey 7-S model was derived from Leavitt's work.

As stated previously, an organizational model can be useful in a number of ways:

- It can help to categorize and summarize.
- It can help to enhance understanding. For example, finding that problems exist in four categories but not in six others helps to establish priorities for action.
- It can help to interpret data about the organization. In other words, how categories in the model are arranged vis-à-vis one another can aid in determining what action to take regarding change. For example, in Weisbord's six-box model, the leadership box (category) is central, with the remaining five boxes surrounding leadership. For Weisbord, leadership serves as the coordinating mechanism for the other five categories, attempting to keep them in balance. In Weisbord's thinking, for any action to be taken toward change, leadership will inevitably play a key role.
- It can help to provide a common, shorthand language. Simply using the category culture, for example, covers a lot of ground.
- It can help to guide action for change. In other words, it can serve as a "cognitive map." If some of the categories are more important than others with regard to change, how the categories are situated in the model vis-à-vis one another can provide direction for change that is likely to be clearer than otherwise might be the case.

As previously mentioned, an organizational model can indeed be useful, but there are at least two important caveats:

Content is critical: A model is only as good as the components selected and the arrangement of these components—that is, how each part relates to the others. Organizations consist of many components. The myriad parts need to be categorized into a workable number and it must be determined which components are the most important.

Metaphors are paradoxical: In 1997, Gareth Morgan stated this caveat eloquently: "Metaphor [a model is a metaphor after all] is inherently paradoxical. It can create powerful insights that also become distortions as the way of seeing created through a metaphor becomes a way of not seeing" (p. 5).

7. FOUR MODELS COMMONLY USED BY OD PRACTITIONERS

7.1. Weisbord's Six-Box Model

The six boxes are leadership (which is central) surrounded by purposes, structure, rewards, helpful mechanisms, and relationships. The six are arranged in a circle surrounded by environment as a seventh variable. The primary diagnostic question that Weisbord asks for each box helps to define that particular category. The questions are as follows:

- Leadership: Does someone keep the boxes in balance?
- Purposes: What business are we in?
- Structure: How do we divide the work?
- Rewards: Do all needed tasks have incentives?
- Helpful mechanisms: Do we have adequate coordinating technologies?
- Relationships: How do we manage conflict among people? With technologies?

Weisbord's model is useful when time is limited (i.e., when a simple framework is needed for quick service) and/or when the client is not particularly sophisticated regarding systems thinking.

7.2. Nadler–Tushman Congruence Model

This is a more sophisticated model but it is based on the assumption that Weisbord *et al*. posit that an organization is an open system influenced by its external environment. The primary inputs from the environment are environmental forces, such as competition, resources (e.g., raw materials and/or capital), and history (the organization's past). Strategy then connects the inputs to the four throughputs: informal organization (e.g., culture), formal organizational arrangements (structure and accountability systems), individual (the people), and task (nature of the work). These throughputs lead to outputs, which amount to performance at three levels: organizational, group, and individual. The primary premise of the model is the importance of congruence, the concept of fit. Therefore, this model goes beyond mere description and takes the position that an organization's effectiveness is a function of how well the components of the model fit together. For example, is strategy congruent (i.e., responding well) to forces in the external environment? Do employees' skills and abilities match well with the tasks of the organization? In summary, the components of the Nadler–Tushman model are fairly comprehensive and have good face validity.

7.3. Tichy's TPC Framework

This model is much more explicit about organization change and states that the nine components of the model are "change levers": (i) external interface, or the organization's external environment; (ii) mission; (iii) strategy; (iv) managing organizational mission/strategy processes (i.e., realistically engaging the relevant interest groups); (v) task—change often requires new tasks; (vi) prescribed networks—more or less, the formal organizational structure; (vii) organizational processes—communicating, problem solving, and decision making; (viii) people; and (ix) emergent networks—more or less, the informal organization.

TPC stands for technical, political, and cultural. These three "systems" overlay or cut across the nine levers, forming a matrix consisting of subcategories of, for example, the technical system with mission and strategy or the political system with people, etc. The technical system represents rationality and is strongly associated with empiricism and the scientific method. With the political system, Tichy argues that organizations have dominant groups with bargaining being the primary mode of change. The cultural system is the organization's norms and values as well as shared symbols. Although Tichy also believes that congruence among all these elements is important, he prefers the term alignment. Data are collected and then

categorized within the matrix of the three systems with the nine levers.

7.4. Porras's Stream Analysis

Porras developed a framework rooted in open system theory, as are the previously discussed models. His input, of course, is the external environment and the output is defined as both organizational performance and individual development, a unique aspect not included in other models. The throughput consists of two broad categories: organizational members' behavior (individual cognitions and on-the-job behavior) and work setting or context. The work setting is composed of four components that Porras labels as streams: (i) organizing arrangements (goals, strategies, structure, and systems), (ii) social factors (culture, networks, and individual attributes), (iii) technology (equipment, machinery, job design, etc.), and (iv) physical settings (space, interior design, ambiance, etc.). In a sense, Porras's model also concerns congruence. His diagnostic process emphasizes a pictorial chart of how the four main streams flow vis-à-vis one another. His stream diagrams remind one of flowcharts designed by an engineer.

7.5. Burke–Litwin Causal Model of Organization Performance and Change

This model is also about theory. Litwin and Burke make the argument that certain organizational variables have more "weight" than others and therefore tend to have effects on less powerful organizational variables. However, the arrows in the model go in both directions and therefore causes and effects are bilateral. Nevertheless, some variables are more powerful than others in causing change. The model is intentionally arranged in a hierarchical manner, with the top boxes referred to as transformational factors (Fig. 4). When practitioners are working to change these categories—mission/strategy, leadership, and culture—the organizational effect is more revolutionary than evolutionary, more discontinuous and episodic than continuous. On the other hand, when practitioners are working to change the categories in the lower remaining part of the model, referred to as the transactional factors, the organizational effect is more continuous than episodic, etc. (Fig. 5).

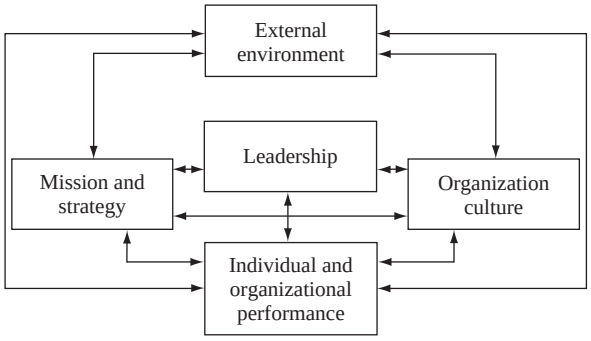


FIGURE 4 The transformational factors. From the Burke–Litwin model of organizational performance and change. © 1987, 1992 by W. Warner Burke, Associates, Inc.

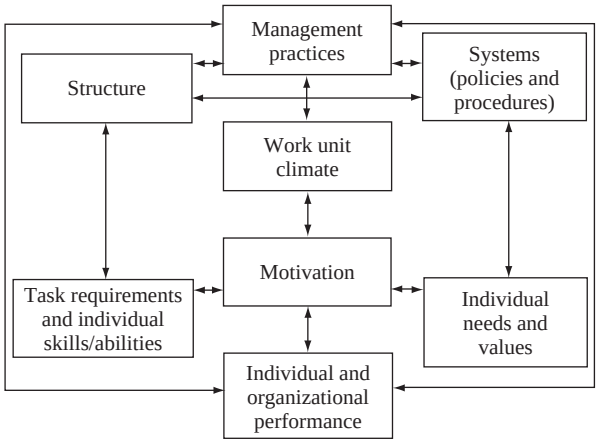


FIGURE 5 The transactional factors. From the Burke–Litwin model of organizational performance and change. © 1987, 1992 by W. Warner Burke, Associates, Inc.

When compared with the transformational categories, the transactional factors represent activities and organizational dimensions that are concerned more with daily operations, are more incremental with respect to change, and are more related to foreground (e.g., work unit climate) than to background (e.g., culture).

In summary, the five models discussed in the previous sections are the most prominent ones in the OD literature, but they may not be the ones used most. Many OD practitioners construct and develop their own models. Regardless of which model is used, more experienced OD practitioners rely on one that is grounded in open system theory and deployed in a manner consistent with action research.

8. PLANNING AND MANAGING THE CONSEQUENCES OF CHANGE

It is appropriate to plan change in terms of phases and steps, and it is appropriate to use models that help to frame data and to plan further. However, the actual process of organization change does not occur in the orderly way suggested in this article. This section addresses the realities of organization change.

The paradox of planned organization change is that the change is planned in a linear fashion—that is, phase 1 followed by phase 2, stage 1, 2, 3, etc., or steps 1–7; however, when the change is actually implemented, the process becomes nonlinear. As stated previously, the carrying out of change in an organization is messy: Things do not proceed exactly as planned; people do things their own way, not always according to the plan; some people resist or even sabotage the process; and some people who would be predicted to support or resist the plan actually behave in the opposite way. In short, unanticipated consequences occur. Leaders of change often say something like, "For every step forward we take, we seem to fall back two steps; something always needs fixing to get us back on track."

Provided the change goals are clear and change leaders are willing to stay the course, over time the process may become somewhat linear, or at least a pattern may emerge. However, linearity is not what anyone experiences during the implementation process, which may seem chaotic, with people in the organization constantly asking "Who's in charge here?"

With respect to managing the process, change leaders may be operating according to plan, especially with the change goal(s) clearly in mind, but the bulk of their time and energy is spent on dealing with issues and problems that were not planned for, were not anticipated, and must now be addressed and resolved. Otherwise, the change effort will become stymied, people will become frustrated if not disillusioned and eventually cynical, and ultimately true organization change will have failed. If planned and managed well, OD efforts can succeed.

9. DOES OD WORK?

This is the basic question that most managers in organizations ask when considering a proposal for getting involved in an OD effort. The answer to the question is "yes," provided the principles and practices summarized in this article are applied.

With respect to research, some representative and key studies may be considered. In their study of 11 different types of OD initiatives employing a meta-analysis (a statistical procedure for comparing results of a number of studies that investigate related content) of 98 studies, Guzzo *et al*. concluded that "psychologically based organizational interventions," as they called them, increased worker productivity. Some of the initiatives or interventions studied were job redesign, participative management, team building, and socio-technical systems methods. These interventions are common in OD practice. Also common in evaluating whether OD works are such measures of performance, broadly defined, as worker productivity, employee attitudes, job satisfaction, financial results (e.g., profits and/or cost reduction), customer satisfaction, turnover and absenteeism, quality of output, and degree of innovation.

In another study using a meta-analysis approach of 126 studies, Neuman *et al*. found that OD interventions had a significantly positive effect on worker attitudes and job satisfaction. In addition, their study showed that multiple OD interventions (e.g., intervening simultaneously with team building, leadership development, and an organizational redesign) had greater effect for positive change than relying solely on one OD intervention. In a less rigorous study of seven successful OD efforts, Burke found that the common denominator across all seven was the fact that multiple interventions were used in each case.

9.1. Evidence

There is evidence that OD works. However, there are problems concerning this body of work when matters of definition, purpose, and methodology are considered.

When discussing OD, what exactly are we talking about? An intervention or two or a total system change? Different OD interventions may result in different outcomes. With respect to purpose, is the study for knowledge generation or to evaluate the effectiveness of OD?

9.2. Methodology

It is not easy to conduct OD research using "normal" or traditional scientific methods. There are important considerations to be made when, for example, there is no control group and when it is not clear which variables of study are independent and which are dependent. Also, many studies have shown that the

researcher and methods used can have powerful consequences (biases) on outcomes.

Although there are complex issues to deal with regarding the determination of the effectiveness of OD, there seems to be cumulative evidence that, on the whole, OD works. Who makes it work? It is, of course, the managers and work unit members in the organization. The OD consultant plays a critical role as well.

9.3. The OD Consultant

Often, the term OD practitioner is used instead of OD consultant; in either case, the OD practitioner functions as a consultant. The OD consultant may be either internal (a full-time or part-time employee) or external (a member of a consulting firm, self-employed, or an academic who consults part-time). Although there are differences (e.g., it is difficult for an internal OD person to be a consultant to the CEO of his or her organization and to be as objective about issues in the organization as may be desirable, and the external consultant will never fully understand the culture of the client organization without living in it), the two still operate very similarly regarding the phases of OD practice (entry, contracting, diagnosis, etc.). Ideally, the internal consultant works in tandem with an external OD consultant.

With respect to the role of an OD consultant, Schein's distinctions are clarifying. He described three models or approaches: the purchase model, the doctor–patient model, and the OD model. The purchase model consists of expert services and information that a client organization may buy. For example, a client organization may want a market research study to be conducted and will purchase the services of a firm specializing in this form of expertise. The doctor–patient model means that the client organization calls in the consultant and describes a problem that needs attention, such as a recent loss of market share, an outdated information technology system, or the loss of too many talented people. The client then expects the consultant to provide a remedy. The OD consultant is more facilitative (i.e., helps the client organization to solve its own problems). Schein (1987) describes the process clearly:

> It is a key assumption of change that the client must share in the process of diagnosing what may be wrong (or learn to see the problem for himself), and must be actively involved in the process of generating a remedy because only the client ultimately knows what is possible and what will work in his culture and situation. (p. 30)

The primary but not exclusive function of the OD consultant is therefore to help clients learn how to help themselves more effectively. The OD consultant may provide expertise, and even prescribe a remedy, but the role is most often one of being a facilitator (e.g., helping an individual manager to gain more self-insight about his or her impact on others, helping a work unit to operate more as a team, or helping the entire organization to understand the results of a systemwide survey).

9.4. Expert Facilitation

Expert facilitation requires certain individual skills and abilities. At least 10 are key:

1. Tolerance of ambiguity: Organizations are not neat and orderly and neither are the data that the OD consultant collects and has to analyze and report to the client.
2. Influence skills: The OD consultant's authority comes from knowledge, of course, but equally as important is the ability to impart that knowledge in a way that makes a difference but is not perceived as arbitrary.
3. Confronting difficult issues: Much of OD practice is bringing to the surface matters that clients are not eager to face.
4. Ability to support and nurture others: This skill is especially critical during times of conflict and stress.
5. Listening well and with empathy: This is particularly important for interviewing and when the client is under stress.
6. Being able to recognize one's feelings and intuitions quickly: This is especially important to discern between one's own perceptions and feelings and those of the client.
7. Conceptualization: This involves being able to express in clear language possible linkages in the organization, such as cause–effect and if–then combinations, and within a systems context.
8. Ability to discover and mobilize human energy: This includes within oneself and the client organization. One should remember that there is energy in resistance; channeling that energy is key.
9. Being a teacher: There are occasions when the OD consultant needs to give a "5-minute lecture" about knowledge that relates directly to the problem at hand.
10. Maintaining a sense of humor: The consultant and/or the client taking themselves too seriously deflects thoughts and energy that could otherwise be

applied more usefully. Also, humor can help to reduce tension when timed well.

This list is a compilation of highly important abilities and skills for the OD consultant. Other abilities could be itemized; this list is a starting point.

9.5. Becoming an OD Consultant

Becoming an OD consultant can be considered as a process of following three paths—apprentice, academic education, and professional development. First, there is no substitute for experience. Being an apprentice to someone with considerable OD experience is perhaps the most useful way to learn OD. After all, OD consultants are also known as practitioners, and the more practice the better. Practice alone is not enough, however. Thus, serving as an apprentice is valuable since one can receive feedback on his or her practice from a more experienced practitioner in the field.

Second, regarding academic education, many universities provide a curriculum in OD, some with master's or even doctoral degree programs in OD or in related fields. The following subjects are suggested as a guide for courses to take that would be the most useful:

- A basic, survey course in organizational psychology or organizational behavior
- Group dynamics
- Research methods that include both quantitative and qualitative skills—understanding and analyzing data are critical to effective OD practice
- Adult learning, especially how to learn from experience
- Career development
- Counseling and coaching, including interviewing skills
- Organization development and change—an overview course of the field
- Training and development—designing programs and other learning activities
- Action research
- Human resource management—often the OD function in an organization is a part of the larger human resource or personnel function
- Consulting skills with emphasis on process consultation
- Organizational dynamics and theory—how organizations work (and do not work) and a comparison of different types of organizations
- Functions of organization—learning the "business" of organizations (i.e., marketing, accounting and finance, strategy, operations, information, technology, budgeting, etc.)

Third, regarding nonacademic training or professional development, the following are suggestions:

- Attend a laboratory, experiential training program to gain greater self-awareness.
- Attend workshops on how to consult; often these are activities that precede large conferences.
- Attend conferences that focus on OD or related practices (e.g., participating in the OD network's annual meeting and regional events).
- Attend specific professional development programs, such as the Principles and Practices of OD and the Advanced Program in OD and Human Resource Management sponsored by Teachers College, Columbia University, or the organization and systems development programs offered by the Gestalt Institute of Cleveland.

Finally, as noted previously, being an OD consultant means being a practitioner. Combining content learning from academic and professional development activities with practice and practice coupled with feedback on this practice are the necessary ingredients for learning and becoming an OD consultant.

10. OD VALUES AND CONSIDERATIONS FOR THE FUTURE OF THE FIELD

10.1. OD Values

Organization development emerged from beliefs that people in organizations had much of the knowledge required to make the system more effective. It was a matter of eliciting and applying this knowledge. This way of believing and acting was steeped in, if not dependent on, antecedent beliefs in and respect for human dignity, diversity, openness in communication, and the importance of empowering organizational members to take initiative and embrace change. Thus, OD grew from humanistic beliefs. However, there have always been the needs and demands of the organization. Thus, organizational requirements, particularly meeting budget demands and "bottom-line" pressures, are often seen as conflicting with humanistic values. Often, OD consultants spend more time and effort on the "D" and less on the "O," but both are equally important. Weisbond (1977) addressed this potential conflict when he stated

> OD's right goal—its central purpose—grows from its proper setting. If the proper setting is organizations, then there is only one right goal for OD: to confront an issue which most certainly predates the industrial revolution. That is the tension between freedom and constraint. OD's right purpose is to redress the balance between freedom and constraint. (p. 4)

There is always tension between the autonomy of the organizational member and the organization's requirements. The primary mission of the OD consultant, therefore, is to redress the balance whenever either one is out of kilter.

10.2. Considerations for the Future of OD

The field of organizational development originated during the late 1950s and early 1960s. OD today is well established. It now seems, however, that OD has reached a rather stagnant state. It is as if all that could be gleaned from the behavioral sciences to apply to organization change and development has been. Undoubtedly this is not true, but in any case, much of the creativity happened in the 1960s and early 1970s, with little innovation since then. Exceptions to this statement include the emergence of appreciative inquiry, self-directed groups, and large group interventions; otherwise, OD appears to be resting on precedence. New ways of thinking about organizations and changing them are needed. Within the field of OD, deeper understandings of the impact of the external environment on organizations and the difficulty of changing organizational culture, executive leadership, and organizational structure are needed. Perhaps of more importance is to infuse new ideas and theory from outside OD that are applicable and relevant nevertheless. An example is theory from the life sciences. The growing knowledge from chaos theory and nonlinear complex systems theory may be nascent at present but brimming with ideas for how to think differently about organizational behavior and change. Finally, even from the more popular literature, a book such as that by Gladwell, titled *The Tipping Point*, is loaded with ideas directly applicable to OD.

For OD to survive and thrive in the future, an infusion of new and different thinking is required. The value of continuous or lifelong learning is absolutely necessary, and much of this learning needs to be done outside the behavioral sciences.

See Also the Following Articles

Environmental Psychology, Overview ■ Industrial/Organizational Psychology, Overview ■ Organization Development ■ Work Motivation

Further Reading

Bunker, B. B., & Alban, B. T. (1997). *Large group interventions: Engaging the whole system for rapid change*. San Francisco: Jossey-Bass.

Burke, W. W. (2001). The broad band of organization development and change: An introduction. In L. Carter, D. Giver, & M. Goldsmith (Eds.), *Best practices in organization development and change* (pp. 3–9). San Francisco: Jossey-Bass/Pfeiffer.

Burke, W. W. (2002). *Organization change: Theory and practice*. Thousand Oaks, CA: Sage.

Burke, W. W. (2004). Some key examples of knowledge management. In M. Goldsmith (Ed.), *Leading organizational learning* (pp. 269–280). San Francisco: Jossey-Bass.

Burke, W. W., & Noumair, D. A. (2002). The role and function of personality assessment in organization development. In J. Waclawski, & A. H. Church (Eds.), *Organization development: Data-driven methods for change* (pp. 55–77). San Francisco: Jossey-Bass.

Capra, F. (1996). *The web of life*. New York: Anchor.

Church, A. H., & Waclawski, J. (1998). *Designing and using organizational surveys: A seven-step process*. San Francisco: Jossey-Bass.

Cummings, T. G., & Worley, C. G. (2005). *Organization development and change* (8th ed.). Mason, OH: South-Western.

Foster, R. N., & Kaplan, S. (2001). *Creative destruction: Why companies that are built to last underperform the market—and how to successfully transform them*. New York: Currency.

French, W. L., & Bell, C. H., Jr. (1995). *Organization development: Behavioral science interventions for organization improvement* (5th ed.). Englewood Cliffs, NJ: Prentice Hall.

Gladwell, M. (2000). *The tipping point: How little things can make a big difference*. Boston: Little, Brown.

Hultman, K. (2002). *Balancing individual and organizational values*. San Francisco: Jossey-Bass/Pfieffer.

Kraut, A. I. (Ed.). (1996). *Organizational surveys: Tools for assessment and change*. San Francisco: Jossey-Bass.

Morgan, G. (1997). *Images of organizations* (2nd ed.). Thousand Oaks, CA: Sage.

Pascale, R. T., Milleman, M., & Gioga, L. (2000). *Surfing the edge of chaos: The laws of nature and the new laws of business*. New York: Crown Business.

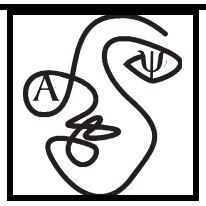
Schein, E. H. (1987). *Process consultation: Vol. 2. Its role in organization development* (2nd ed.). Reading, MA: Addison-Wesley.

Schein, E. H., DeLisi, P. S., Kampas, P. J., & Sonduck, M. M. (2003). *DEC is dead, long live DEC: The lasting legacy of Digital Equipment Corporation*. San Francisco: Berrett Koehler.

Trist, E. (1993). Introduction to Volume II. In E. Trist, & H. Murray (Eds.), *The social engagement of social science. A Tavistock anthology: Vol. 2. The socio-technical perspective* (pp. 36–60). Philadelphia: University of Pennsylvania Press.

Weisbord, M. (1987). *Productive workplaces: Organizing and managing for dignity, meaning, and community*. San Francisco: Jossey-Bass.

Zaccaro, S. J. (2001). *The nature of executive leadership: A conceptual and empirical analysis of success*. Washington, DC: American Psychological Association.



Organizations, Careers in

Ayse Karaevli and Douglas T. Hall
Boston University, Boston, Massachusetts, USA

1. Organizational Careers
2. The New Career Contract: Protean and Boundaryless Careers
3. Predicting Career Effectiveness
4. Conclusion
Further Reading

GLOSSARY

adaptability The capacity and motivation to adapt or change in response to changing task or environmental circumstances.

boundaryless career A career not bounded, not tied to a single organization, not represented by an orderly sequence, and marked by less vertical coordination and stability than an organizational career.

career In both the popular and behavioral science literature, there are four distinct meanings of the term career: career as advancement, career as profession, career as a lifelong sequence of jobs, and career as a lifelong sequence of role-related experiences. Based on all these different interpretations, career can be defined as the individually perceived sequence of attitudes and behaviors associated with work-related experiences and activities over the span of a person's life.

identity The person's awareness of his or her values, interests, abilities, and plans; the degree of integration between past, present, and future concepts of self; and the extent to which the person feels integrated over time.

organizational career A career conceived to unfold in a single employment setting.

protean career A career based on self-direction in the pursuit of psychological success in one's work.

psychological success The feeling of pride and personal accomplishment that comes from knowing that one has done one's "personal best."

This article discusses the general developmental stages that people pass through during the course of their working lives in organizations. The ways in which careers have changed since the mid-1970s are also discussed, with a focus on the changing career contract and the protean, or self-directed, career. It also examines what happens after a career choice has been made, particularly how the effectiveness of a career can be predicted. Performance, career attitudes, adaptability, and identity are discussed as the main dimensions of career effectiveness.

1. ORGANIZATIONAL CAREERS

1.1. Career and Life Stages

People go through a regular series of life experiences and tasks that constitute general phases of human development. Although in an era of rapid change and complexity the concept of regular and predictable stages or phases has diminished in significance, we have not lost completely the experience of regular, predictable changes in our roles, expectations, concerns, and self-images that are packaged as life and career stages. There are several life and career models that explain these regular stages or phases.

©2004 Elsevier Inc.
All rights reserved.

In Erik Erikson's 1963 theory of life stages, there are eight stages of development, each of which is characterized by a particular developmental task that the person must work through before advancing fully to the following stage. In the realm of vocational behavior, Donald Super and associates employed a model of five developmental stages: childhood, adolescence, young adulthood, maturity, and old age. According to their model, in childhood the main task is growth. In the adolescence stage, a person explores his or her interests and different specific career opportunities. In adulthood, a person may initially flounder a bit but eventually establish himself or herself in a particular field. In maturity, when the generativity concerns would probably be the most important, the person continues to hold his or her own in a sort of career plateau. Old age is a period of disengagement when ego integrity is achieved if the person has resolved all the earlier stages.

Psychologist Daniel Levinson and colleagues developed a theory of life stages and transitions based on in-depth clinical case studies of a sample of men and women. They utilized the concept of life structure, which is formed by the individual's sociocultural world, aspects of the self that are played out and those that are inhibited or neglected, and the person's participation in the world. In 1968, Hall and Nougaim identified organizational career stages by using a three-stage model of career development: establishment, advancement, and maintenance. At the establishment stage, people have strong concerns about safety, gaining recognition, and establishing themselves in the organization. The advancement stage is characterized by concerns for promotion and achievement. The maintenance stage represents the onset of a plateau, where the incumbent has cues that he or she is nearing the limit of his or her advancement.

1.2. Schein's Model of the Organizational Career

A model of career development in organizations needs to describe the career from two separate but related perspectives: the career as described by the characteristics and experiences of the person, who moves through the organization, and the career as defined by the organization, which involves policies and expectations. A useful approach for understanding this relationship was proposed by Edgar Schein in 1971, who viewed the organization as a three-dimensional space like a cone, in which the three dimensions represent

- Vertical: Moving up or down, representing one's changing rank or level in the organization
- Radial: Moving more (or less) "inside" in the system, becoming more (or less) central, part of the "inner circle" acquiring increased (or decreased) influence in the system
- Circumferential: Transferring laterally to a different function, program, or product in the organization

1.3. Careers as Learning Cycles

One of the keys to understanding the new employment contract is the fact that the employees' needs and career concerns change during the course of the career in a much more dynamic way than in the past. Therefore, in Douglas Hall's 1996 concept of the flexible career or protean career, the career is not measured by age-linked life stages but, rather, by a series of short (3–5 year) learning cycles. Each learning cycle has the following phases: exploration, trial of new activity, mastery, and then exploration of some other new activity. Because the life cycle of technologies and products is short, so too are personal mastery cycles. As a result, people's careers will increasingly become a succession of "ministages" (or short-cycle learning stages) of exploration–trial–mastery–exit as they move in and out of various product areas, technologies, functions, organizations, and other work environments.

2. THE NEW CAREER CONTRACT: PROTEAN AND BOUNDARYLESS CAREERS

2.1. The Protean Career Contract

Another way to frame the changes in career patterns is in terms of the psychological contract between the employee and the organization. Ian MacNeil discussed two forms of what he called the social contract. The first was relational, which was based on assumptions of a long-term relationship and trust that the relationship would be a mutual one. In contrast, the other form was transactional, based on a shorter term exchange of benefits and contributions. MacNeil was discussing the role of an individual in a larger society, but his concepts seem applicable to organizations as well.

From the perspective of the individual, there is a shift from the organizational career to what Hall described as the protean career—a career based on self-direction in the pursuit of psychological success in one's work. The driving forces in the protean career are the person's core values and need for control of his or her life and career choices. In addition, the protean person has a strong learning or exploratory orientation that would predispose him or her to be quite comfortable in a career characterized by continuous learning cycles, as described previously.

The measure of success for the protean careerist is psychological success, or success as measured by the person's own goals and values. Thus, in Everett Hughes' terms, to understand the protean career, the observer must examine the subjective career, which is the way the individual experiences his or her career. In contrast, the more traditional organizational career was often assessed from the perspective of the objective career, which is the way career success would be judged by an external observer, usually in terms of outcomes such as money, power, or position.

2.2. The Boundaryless Career

In 1996, Michael Arthur and Denise Rousseau identified a related concept, the boundaryless career. They defined the boundaryless career as not bounded, not tied to a single organization, not represented by an orderly sequence, and marked by less vertical coordination and stability. In their words, "boundaryless careers are the opposite of 'organizational careers'—careers conceived to unfold in a single employment setting" (p. 5). In 1993, Defillipi and Arthur offered a congruent explanation of the boundaryless career that evolved around interfirm mobility. Peiperl and Baruch, in their discussion of the "post-corporate career," agreed that the emerging career clearly centered outside the organization—increasingly identifiable with professional and familial as opposed to organizational communities.

Thus, if the old contract was with the organization, in the protean career the contract is with the self and one's work. In a study comparing data collected in 1978 and 1989, Linda Stroh and colleagues found evidence of this shift from an organizational focus to an investment in one's own work; satisfaction with the company decreased from 1978 to 1989, but job involvement and job satisfaction increased. (They also reported that managers in 1989 were changing jobs and relocating more often than those in 1978.)

2.3. Mobility via Contingent Employment

One of the ways that career mobility and protean careers are currently manifest is in the rise of contingent employment. In the United States, the largest employer at the time of this writing is a temporary agency, Manpower, Inc. More than 90% of U.S. employers use temporary workers as part of their staffing mix, as identified by von Hipple *et al.* in 1997, and nearly 3 million people worked as temporary workers in a single year. It is clear that one of the major ways that careers have become boundaryless has been through the practice of temporary employment.

2.4. Life and Work Roles

Another way that change and flexibility manifest themselves in contemporary careers is in the interplay between the person's work role and the roles in his or her private life. There are three common forms of stress between work and home, and it is important to consider each one separately. (There is a huge literature on work–life issues, so this discussion is simplified by considering married couples.) The first and most common source is role overload. This is caused by the sheer number and intensity of demands on a couples' time. Overload pressure is especially high for young couples (at a career stage in which financial resources are tight) with children. The sheer number of demands on the partners exceeds the time and energy to do them.

The second type of work–life stress derives from role conflict, the incompatibility of different role expectations. One type of conflict is interrole conflict, in which one partner's different roles make mutually conflicting demands, such as when a person, as an employee, is expected to be out of town for a client meeting and is also expected, as a parent, to be at his or her child's school for the fifth-grade play. It is not possible to be in both places at the same time (although people may be tempted to do so electronically). There is also intersender conflict, within a role. An example is when a father is home taking care of his daughter and she wants him to let her stay up past her bedtime to watch a special TV show, even though the mother had left strict instructions that the daughter should be in bed on time.

A third type of conflict is intrapersonal conflict (within the person), in which a person feels torn between two valued activities (e.g., going to the son's band concert or to the daughter's all-star softball game). It could also be a conflict of unmet expectations, in which the person

feels that he or she is simply not measuring up to the standards set by other people he or she respects ("I could never be the father that he is" or "I'll never be as successful as she is").

The third source of stress is change. This is especially problematic for dual-career couples, who may be more likely to experience relocations more often than one-career families. A relocation can trigger all sorts of other changes—new jobs, new home, new friends, new schools, new community, new culture, often a more expensive house and the concomitant financial stress, etc. Again, synchronicity is an issue: The more roles that are changing simultaneously, the more the stress is compounded.

3. PREDICTING CAREER EFFECTIVENESS

In the research literature, career effectiveness has been defined in terms of performance and the popular symbols of success—money and position. However, there are other important measures of career effectiveness: career attitudes, adaptability, and identity. These four dimensions of career effectiveness can be considered as representing a 2×2 grid, as shown in Table I, where one axis involves the time span (short term vs long term) and the other involves the focus (task vs personal) aspect of career development.

3.1. Performance

The major category of performance effectiveness is related to typical financial indicators, such as current salary, average yearly salary increases, or salary in relation to that of other people with equivalent length of service or of similar age. The other major category involves the position, which has been employed as rank or level in an organizational hierarchy, number of promotions received over a given time period (i.e., rate of advancement), or, conversely, length of time in current position.

TABLE I
Four Types of Career Effectiveness[a]

	Time span	
Focus	*Short term*	*Long term*
Task	Performance	Adaptability
Self	Attitudes	Identity

[a]Adapted from Hall (2002).

3.2. Career Attitudes

The way the career is perceived and evaluated by the individual defines an individual's attitudes toward his or her career. The major attitudes related to career are organizational identification and commitment, which tap the psychological involvement that the person feels with the organization and the extent to which his or her organizational membership is a significant aspect of his or her identity; the involvement and commitment to career and family and negative experiences related to involvement and commitment, such as burnout; the psychological success, which is the feeling of pride and personal accomplishment that comes from knowing that one has done one's personal best, as Hall and Mirvis discussed; and self-efficacy, which is the individual's beliefs about his or her ability to perform successfully a given task, as identified by Bandura.

3.3. Adaptability

Morrison and Hall define adaptability as the capacity and motivation to adapt or change in response to changing task or environmental circumstances. Hall proposed adaptability as a higher order quality called metacompetency—the capacity to master many more specific skills when one masters this metacompetency.

3.4. Identity

An individual's sense of identity can be defined as (i) the person's awareness of his or her values, interests, abilities, and plans; (ii) the degree of integration between past, present, and future concepts of self; and (iii) the extent to which the person feels integrated over time. Role transitions, socialization, life stage, as well as race, culture, and history are critical factors that facilitate the growth of the person's identity. There are two facets of the person's overall sense of self. Personal identity refers to those qualities in the person's self-perceptions that are unique to him or her. Social identity descries the portion of the overall identity that derives from the person's membership in a particular social group, such as a gender group or a racial or ethnic group.

4. CONCLUSION

In the postindustrial era, individual career effectiveness entails more than traditional "objective" indicators of

success, such as money and position. In a knowledge economy, the rapid rate of change puts positive attitudes, identity growth, and adaptability at a premium. These are all qualities that are more "subjective," or internal and unique to the person. The ultimate evaluation of these career outcomes (i.e., career success) can be assessed from the individual's perspective. This individual or subjective definition of success is called psychological success.

All these personal career outcomes, including psychological success, are developable—that is, they can be influenced by the individual and the organization. Also, they tend to go together and to reinforce one another. For example, people with positive attitudes tend to make jobs into satisfying and developmental experiences; that is, they adapt effectively to frustrations in the work environment. Knowing oneself enables the person to make better choices that are in line with the person's interests and values and to know whether it is better to make adaptations or to move on and select a new place to work. In turn, effective adaptation creates new skills and relationships, which in turn lead to a more positive and proactive self-identity.

Acknowledgment

Portions of this article are based on and adapted from Hall (2002).

See Also the Following Articles

Career Counseling ■ Holland's Theory (Vocational Personality Types) ■ Indecision, Vocational ■ Vocational Assessment in Schools ■ Vocational Interests ■ Vocational Psychology, Overview ■ Work Role, Values Sought in the

Further Reading

Arthur, M. M., & Rousseau, D. M. (Eds.). (1966). *The boundaryless career: A new employment principle for a new organizational era*. New York: Oxford University Press.

Hall, D. T. (1996). Protean careers of the 21st century. *Academy of Management Executive*, *10*(4), 8–16.

Hall, D. T. (2002). *Careers in and out of organizations*. Thousand Oaks, CA: Sage.

Hall, D. T., & Mirvis, P. H. (1996). The new protean career: Psychological success and the path with a heart. In D. T. Hall *et al*. (Eds.), *The career is dead—Long live the career: A relational approach to careers* (pp. 15–45). San Francisco: Jossey-Bass.

Levinson, D. J. (1986). A conception of adult development. *American Psychologist*, *41*, 3–13.

Levinson, D. J. (with Levinson, J.). (1996). *The seasons of a woman's life*. New York: Knopf.

Levinson, D. J., Darrow, C. N., Klein, E. B., Levinson, M. H., & McKee, B. (1978). *The season's of a man's life*. New York: Knopf.

MacNeil, I. R. (1980). *The new social contract*. New Haven, CT: Yale University Press.

Peiperl, M., & Baruch, Y. (1997, Spring). Back to square zero: The post-corporate career. *Organizational Dynamics, 23–34*.

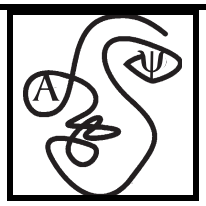
Schein, E. H. (1971). The individual, the organization, and the career: A conceptual scheme. *Journal of Applied Psychology*, *7*, 401–426.

Stroh, L. K., Brett, J. M., & Reilly, A. H. (1994). A decade of change: Managers' attachment to their organizations and their jobs. *Human Resource Management*, *33*, 531–548.

Super, D. E. (1992). Toward a comprehensive theory of career development. In D. H. Montross, & C. J. Shinkman (Eds.), *Career development: The theory and practice* (pp. 35–64). Springfield, IL: Thomas.

Super, D. E., & Bohn, M. J., Jr. (1970). *Occupational psychology*. Belmont, CA: Wadsworth.

Wiese, B. E., Freund, A. M., & Baltes, P. B. (2002). Subjective career success and emotional well-being: Longitudinal predictive power of selection, optimization, and compensation. *Journal of Vocational Behavior*, *60*, 321–335.



Overtraining and Burnout in Sports

Michael Kellmann

Ruhr-University of Bochum, Bochum, Germany

1. Introduction
2. Recovery
3. Overtraining
4. Burnout
5. Interrelation of Stress States and Recovery Demands
6. Treatment of Overtraining and Burnout
Further Reading

GLOSSARY

burnout A psychological, physical, and emotional withdrawal from a formerly enjoyable and motivating activity resulting from prolonged or chronic stress.

long-term overtraining During long-term overtraining, lasting 3 weeks or more, athletes are on a chronic performance plateau that cannot be influenced positively by short amounts of rest and recovery periods.

overtraining This is due to a long-term imbalance between stress and recovery (i.e., too much stress combined with too little regeneration).

recovery An inter- and intraindividual multilevel (e.g., psychological, physiological, and social) process for the reestablishment of performance abilities. Recovery includes an action-oriented component, and self-initiated activities (proactive recovery) can be systematically used to optimize situational conditions to build up and to restore personal resources and buffers.

short-term overtraining Short-term overtraining, lasting less than 3 weeks, is a regular part of athletic training that leads to a state of overreaching in affected athletes. This state of overreaching is characterized by a transient underperformance that is reversible within a short-term recovery period of 1 or 2 weeks and can be rewarded by an increase in performance ability.

This article addresses underrecovery as a key component in the development of overtraining and burnout. A constant lack of recovery or disturbed recovery results in overtraining or burnout, which frequently occur in sports. Even being slightly underrecovered over a longer period of time results in underperformance.

1. INTRODUCTION

At the 2002 Soccer World Championships in Japan and Korea, teams that were favored to win the World Cup, such as France and Portugal, did not survive the first round in the tournament. During and immediately after the World Cup, some journalists and the World Soccer Association (FIFA) started a discussion about why these teams were not able to perform at their expected level. A connection was quickly made to burnout, overtraining, and underrecovery due to the game schedule of the players during the season. Especially athletes from successful clubs playing in international competitions, such as the Champion League or the Uefa-Cub in Europe, have a very tight competition calendar. For example, Hertha BSC played more games than any of the other German professional soccer teams in the first half of the 1999/2000 season—29 games in 18 weeks, which is an

©2004 Elsevier Inc.
All rights reserved.

average of 1 game every 4.3 days, including travel time. By midseason, the players described themselves as physically and mentally drained. In addition to this schedule, some players are drafted to their national teams and have to play for them in addition to playing regular club games. Therefore, a common tendency seems to be that players purposely miss games of their national teams by claiming an injury to reduce the number of games they must play and to get more time off. Interestingly, often it is not the training and competition that are perceived to be stressful by the players but, rather, traveling and the continuous adaptation to new setting and surroundings. The schedules of professional sports teams in North America, such as basketball, baseball, and hockey teams, are similar to or even more intense than those of European soccer teams. In high-performance sports, especially professional sports, the drills performed in practice are often driven by the competition schedule so that there is only time for regenerative practice. For example, National Hockey League (NHL) teams played 82 games during the regular season in 2000/2001. According to the official game schedule, the Colorado Avalanche, like other NHL teams, played an average of 1 game every 2.3 days (including travel time), and reaching the playoffs added more games to the schedule. How much recovery time is left between games? When 2 or 3 games are played per week, physical and mental recovery is often incomplete. However, lack of recovery or underrecovery is not only due to the frequency of competitions. It can also be enforced by training mistakes, such as (i) having a monotonous training program, (ii) more than 3 hours of training per day, (iii) more than a 30% increase in training load each week, (iv) ignoring the training principle of alternating difficult and easy training days or by alternating 2 difficult days followed by an easy training day, (v) no training periodization and respective regeneration microcycles after 2 or 3 weeks of training, or (vi) no rest days.

2. RECOVERY

Often, recovery is defined as the compensation of deficit states of an organism (e.g., fatigue or decrease in performance) and, according to the homeostatic principle, a reestablishment of the initial state. In 2001, Kellmann and Kallus stated that recovery encompasses active processes of reestablishing psychological and physiological resources and states that allow one to tax these resources again. This more precisely describes the complex issue of recovery. However, it is still not sufficient to define recovery as an elimination of fatigue or system restart. Kellmann und Kallus developed a more precise definition:

> Recovery is an inter- and intraindividual multilevel (e.g., psychological, physiological, social) process in time for the reestablishment of performance abilities. Recovery includes an action-oriented component, and those self-initiated activities (proactive recovery) can be systematically used to optimize situational conditions to build up and to refill personal resources and buffers. (p. 22)

Recovery processes in competitive sports support the restoration of individual activity conditions and well-being following training and competition demands. Several studies have shown that recovery is faster when subjects exercise moderately instead of resting passively. Active recovery consists of cool-down activities, such as muscle relaxation and stretching, after practice and competition. The purpose of these exercises is to eliminate the effects of fatigue through physical activity. Passive recovery, on the other hand, includes treatments such as massages, hot and cold baths, steam baths, and sauna baths, which initiate physiological reactions through physiological stimuli (heat, cold, and pressure), affecting blood flow, respiration rate, and muscle tone. Optimal use of available recovery time is imperative for success in sports. Top performances can only be achieved by athletes who can recover quickly during competition and optimally carry out the changes between stress, recovery, and future stress.

3. OVERTRAINING

Kellmann summarized the literature and stated that 60% of females and 64% of males have been overtrained at least once during their running careers. Ninety-one percent of swimmers who were overtrained during their freshmen year in college also experienced overtraining during the following years, whereas only 30% of athletes who were not overtrained in their freshman year had problems later. In addition, overtraining was seen in 33% of Indian national-level basketball players during a 6-week training camp, in 21% of swimmers on the Australian National Team during a 6-month season, and in more than 50% of semiprofessional soccer players after a 5-month competitive season.

Gould and Dieffenbach pointed out the relevance of overtraining in high-performance sports and its importance as a performance-influencing factor identified in

several studies of high-performance athletes. They summarized a study on positive and negative factors influencing U.S. athletes and coaches at the Olympics in Atlanta in which 84 (28%) of 298 athletes reported that they had overtrained for the games and this overtraining had a negative impact on performance. Similarly, 35 of the athletes identified overtraining/not getting enough rest as the primary coaching action that negatively affected their performance. Ten percent of U.S. athletes who participated in the 1998 Nagano Winter Olympics reported that they were overtrained and that this had negative effects on their performance. The need to taper, rest, not to overtrain, travel less, and stay healthy were changes that they would employ if they could prepare again for the Olympics. Hence, these findings demonstrate that both summer and winter Olympic Games athletes identified overtraining as a concern. Gould and Dieffenbach note that the findings should not be assumed to result only from inappropriate physical training. Other factors, such as psychological stress, travel, personality, inadequate rest, type of recovery activity, and sociological issues must be examined in multifaceted models as well.

It is well accepted that overtraining is due to a long-term imbalance between stress and recovery (i.e., too much stress combined with too little regeneration). The term stress includes all training, competition, and additional nontraining stress factors. Social, educational, occupational, economical, nutritional, and travel aspects, time stress, and the monotony of training increase the risk of an overtraining syndrome.

The same training can have different effects on athletes depending on their personal situation, including emotional stressors (e.g., fights with friends or partners and also parents getting divorced), obligations at school, or difficulties in time management (e.g., practice/school/friends). Often, individuals can easily handle these situations, but when a heavy training load is added to an already high "personal package load," the total impact can be too high. Therefore, coaches should be aware of athletes' personal situations and design training accordingly. The most frequent causes of overtraining cited by athletes are (i) too much stress and pressure, (ii) too much practice and physical training, (iii) physical exhaustion and all-over soreness, (iv) boredom due to too much repetition, and (v) poor rest or lack of proper sleep. Therefore, training holism encompasses two ideas: Training must be balanced and varied, and nontraining time has a major influence on training. It is therefore important that all factors outside the realm of the training session be evaluated with regard to their possible negative influence on total fatigue. Insufficient and/or lack of recovery time between practices is the main cause of overtraining syndrome. Factors such as nutrition, sleep deficit, sickness, travel, and competitions increase the negative effect of insufficient recovery.

In 2002, Kellmann discussed the different uses of the term overtraining and related terms in the literature and concluded that the definition by Lehmann *et al.* provides the best integration and description of overtraining. Lehmann and colleagues distinguished between short-term overtraining, which lasts less than 3 weeks, and long-term overtraining that lasts 3 weeks or more. They noted that short-term overtraining (also called overreaching or supercompensation training) is a frequent part of athletic training, which leads to a state of overreaching in affected athletes. This state of overreaching is characterized by a transient underperformance that is reversible within a short-term recovery period of 1 or 2 weeks and can be rewarded by a state of supercompensation (increase in performance ability after 1 or 2 regeneration weeks after a short-term phase of overtraining). Therefore, short-term overtraining or overreaching is a regular part of athletic training and desired by coaches. When overreaching is done inappropriately or done for too long, short-term overtraining turns into long-term overtraining.

During long-term overtraining, athletes are on a chronic performance plateau that cannot be influenced positively by short periods of rest and recovery. Long-term overtraining can be characterized by an inability to train at the customary level. The symptoms include feelings of depression, general apathy, decreased self-esteem, emotional instability, irritability, impaired performance, lack of supercompensation, restlessness, disturbed sleep, weight loss, loss of appetite, increased resting heart rate, increased vulnerability to injuries, and hormonal changes. An important clinical feature of overtraining syndrome is the increased susceptibility to infections with corresponding symptoms, suggesting some kind of impaired immune response. Unfortunately, when athletes realize that performance has stagnated or declined, it is a common tendency to increase the training effort, which leads to an even deeper stage of long-term overtraining.

4. BURNOUT

Although there is no common definition for burnout, there is a consensus that burnout is a reaction to chronic stress. Smith defines burnout as a psychological, physical, and emotional withdrawal from a formerly

enjoyable and motivating activity resulting from prolonged or chronic stress. It is an exhaustive psychophysiological response exhibited as a result of frequent, sometimes extreme, and generally ineffective efforts to meet excessive physical training and psychological demands. It is also important to note that when athletes burnout of sport they do not always discontinue involvement altogether. Some athletes who experience burnout remain active. They believe that they cannot discontinue for various reasons (e.g., family pressure), and this entrapped participation is characterized by lower motivation, emotional detachment, less satisfaction, and poor performance. Burnout results in a decline in sport participation by an athlete as a result of chronic excess stress.

Maslach defines burnout as an individual syndrome including emotions, attitudes, motives, and expectancies. Burnout is a negative individual experience encompassing problems, distress, negative mood states, dysfunction, and negative consequences. Maslach and Jackson characterize the burnout syndrome through emotional exhaustion (i.e., loss of social interest) and reduced personal accomplishment and involvement. In addition to a lack of performance, burnout eventually results in the quitting of prior enjoyable activities. Research on burnout was originally performed in the area of social professions, and it was expanded to other professional groups, such as athletes and coaches. With an increasing level of burnout, the performance of athletes becomes inconsistent and eventually declines. The effect is supported by overtraining, routine, monotony, boredom, and a lack of diversity. This indicates that burnout is caused by a lack of recovery. In the case of an intense schedule during the season and intense training during the off season, a complete physical and psychological recovery seems to be impossible. Therefore, a long recovery phase or a temporary reduction in athletic activities are the only effective approaches to prevent athletes from dropping out of sport. Consequently, coaches and staff need to be sensitive to the recovery demands of athletes.

On the behavioral level, burned out athletes have less input with regard to training, fewer practice days, differences in motivation, poorer coping behavior, and a longer history within the sport. Gould *et al*. showed for tennis players that Smith's model well described the antecedents and consequences of burnout. Thus, one can view a model of chronic stress as a good basis for explaining burnout in athletes. Cohn identified too much practice or competition, lack of enjoyment, and too much pressure from self or others as the main sources of burnout. Furthermore, stress associated with injuries and resumption of training can drastically increase the risk of dropout and burnout.

Burnout in coaches has also been examined. The work environment of coaches involves long hours, demands much mental and emotional energy, and exerts considerable subjective pressure due to standards set by the coaches or by fans and athletes. A survey by Vealey *et al*. demonstrated the importance of the coach's personality as a factor in the development of burnout. The results also supported the contribution of reward factors as buffers against burnout, which can be derived from Smith's model. Perceived rewards and perceived value of the social role, together with excitement, are the most important antiburnout buffers, as evidenced by research in organizations.

In 1996, Kallus *et al*. reported results from a study of 195 semiprofessional and professional coaches. They found (i) differences in the coping strategies of younger coaches as well as between coaches of team and individual sports; (ii) coaches of individual sports who used more cognitive reevaluation strategies, actively and positively coped with stress, and searched for self-affirmation pitied themselves more than coaches from team sports; and (iii) full-time coaches were more burned out and highly stressed, were experiencing less recovery, and were less fit than semiprofessional coaches. In addition, professional coaches showed deficits in coping strategies and therefore may possibly react less effectively in critical situations. It can be concluded that stress in addition to a recovery deficit and a coping deficit leads to a high risk of burnout. More research is needed to fully support the results of Kallus *et al*.; however, these results strongly indicate that recovery is a key variable in the development of burnout.

5. INTERRELATION OF STRESS STATES AND RECOVERY DEMANDS

The previous discussion has shown that recovery is the key factor in preventing overtraining and burnout. In this context, Kellmann proposed a general model that describes the interrelations of stress states and recovery demands (Fig. 1). The basic assumption of this model is that with increasing stress, increased recovery is necessary to remain at the same stress state. However, limited resources (e.g., time) initiate a vicious cycle: Under increased stress and due to the inability to meet increased recovery demands, a person experiences more stress. Recovery demands are defined as

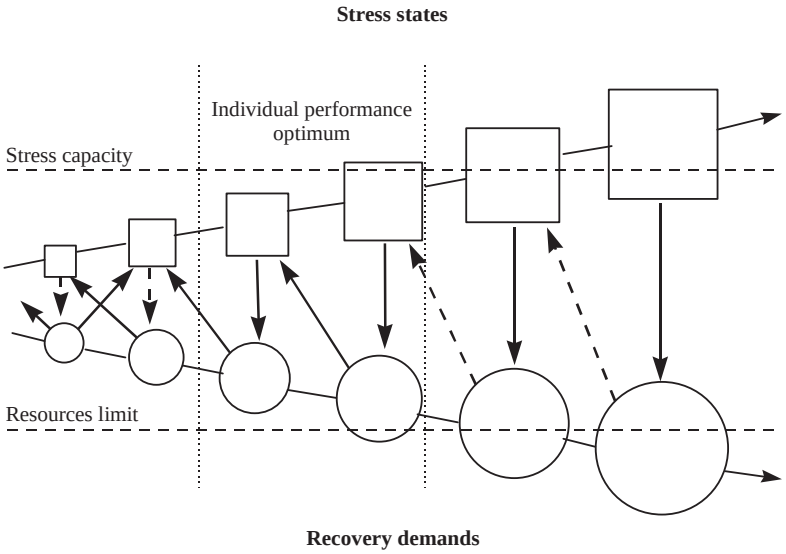


FIGURE 1 The "scissors model" of the interrelation of stress states and recovery demands. Reprinted, by permission, from K. W. Kallus & M. Kellmann, 2000, "Burnout in athletes and coaches" in *Emotions in sport*, edited by Y. L. Hanin (Champaign, IL: Human Kinetics), p. 212.

the quality and/or quantity of recovery activities required to level out the current recovery stress state. People may be stressed to the point that they fail to find or make time to recover adequately or to consider better ways of coping with the situation.

With intermediate levels of stress, one can find an area of optimal performance and thus an area of adequate recovery. Beyond this point, one cannot meet recovery demands without additional recovery activities. Stress will accumulate, and without intervention burnout symptoms are likely to develop. The state of balanced stress and recovery is related to optimal performance. In a state of adequate recovery, the individual can react appropriately and cope successfully with stress without additional recovery activities. A lack of recovery or underrecovery can trigger a process that leads to a state of elevated stress. Because increasing stress limits the possibility of recovery, the athlete must be given opportunities to recover in order to reestablish an optimal level of performance.

With regard to sports, the model may also explains how overtraining develops. The axis of the stress states can be seen as a continuum of an increasing training load, which can be labeled at the end points "no training" and "overtraining." With extended training load, the organismic recovery demands increase proportionally. A short-term planned period of recovery enhances long-term performance effects (e.g., supercompensation). If the training load and intensity increase over a longer time period without adequate or with merely inappropriate recovery, the individual experiences long-term underrecovery, which may result in the overtraining syndrome. To reach the optimal recovery stress state, athletes have to increase their self-initiated activities to fulfill their recovery demands. At each state of the model, recovery can work as a regulation mechanism, which is caused by an increasing distance between the two axes into a longer recovery period (days to weeks). The higher a person is on the stress states or the more extensive the overtraining syndrome, the more recovery efforts are needed to reach the individual optimal recovery stress state. The model interrelation of stress states and recovery demands implies that it is not necessarily bad to be highly stressed as long as a person knows how to recover.

6. TREATMENT OF OVERTRAINING AND BURNOUT

When performance plateaus or instability on performance occur, athletes and coaches have the tendency to increase their effort and the training load, which mostly leads to a vicious cycle. The best treatment for overtraining is rest and recovery phases from 2 to 6 weeks long (sometimes longer), depending on the degree of overtraining. In some cases of overtraining syndrome,

athletes may require months without any training and without any physical activity in order to recover completely. Changes in training can also positively affect the practice routine. A reduction of training load, a change in intensity, or a change in training technique can positively affect athletes. When a coach believes that an athlete is showing signs of overtraining, the athlete should consult a doctor to rule out injuries and damage due to overstrain. If overtraining syndrome is diagnosed, treatment with electrolytes can be considered. Active and proactive recovery is the best treatment for overtraining syndrome. Physical exercises, such as gymnastics, games, regenerative runs, or swimming, as well as a balanced and healthy diet play an important role in the recovery process. Environment and climate changes help athletes to recover both physically and mentally. Activities that were denied because of the rigid training regime can be important too, and psychoregulative techniques, such as progressive muscle relaxation or autogenic training, can accelerate the recovery process.

In the case of burnout, psychoregulative techniques also play an important role, enabling people to deal with chronic stress. Psychoregulative techniques enhance the ability of athletes to have appropriate reaction patterns when needed. Also, when someone is burned out, the setting of short-term and long-term goals is important to gain perspective on a new or formerly enjoyable activity. Setting short-term goals and providing rewards or incentives every time a goal is accomplished is motivating. In addition, having fun goals, especially during the end of a long season, is very beneficial. These goals should be set for competition as well as for practice, especially since practice often comprises more than 90% of the participation time in sports. Gould and colleagues' recommendations for dealing with and preventing burnout include playing for personal reasons, balancing sport with other activities such as school clubs or socializing with friends, stopping participation if the sport is no longer fun to play, focusing on making the sport fun, doing things to relax, and taking time off to recover.

See Also the Following Articles

Performance Slumps in Sport: Prevention and Coping ■ Psychological Skills Training in Sport

Further Reading

Cohn, P. J. (1990). An exploratory study of sources of stress and burnout in youth golf. *The Sport Psychologist*, *4*, 95–106.

Gould, D., & Dieffenbach, K. (2002). Overtraining, underrecovery, and burnout in sport. In M. Kellmann (Ed.), *Enhancing recovery: Preventing underperformance in athletes* (pp. 25–35). Champaign, IL: Human Kinetics.

Gould, D., Tuffey, S., Udry, E., & Loehr, J. (1997). Burnout in competitive junior tennis players: III. Individual differences in the burnout experience. *The Sport Psychologist*, *11*, 257–276.

Kallus, K. W., Kellmann, M., Eberspächer, H., & Hermann, H.-D. (1996). Beanspruchung, Erholung und Streßbewältigung von Trainern im Leistungssport [Strain, recovery, and coping with stress of coaches in elite sports]. *Psychologie und Sport*, *3*, 114–126.

Kellmann, M. (Ed.). (2002). *Enhancing recovery: Preventing underperformance in athletes*. Champaign, IL: Human Kinetics.

Kellmann, M., & Kallus, K. W. (2001). *Recovery–Stress Questionnaire for athletes; User manual*. Champaign, IL: Human Kinetics.

Kreider, R. B., Fry, A. C., & O'Toole, M. L. (Eds.). (1998). *Overtraining in sport*. Champaign, IL: Human Kinetics.

Lehmann, M., Foster, C., Gastmann, U., Keizer, H. A., & Steinacker, J. M. (Eds.). (1999). *Overload, fatigue, performance incompetence, and regeneration in sport*. New York: Plenum.

Maslach, C., & Jackson, S. E. (1986). *Maslach burnout inventory*. Palo Alto, CA: Consulting Psychologists Press.

Smith, R. E. (1986). Toward a cognitive–affective model of athletic burnout. *Journal of Sports Psychology*, *8*, 36–50.

Vealey, R. S., Udry, E. M., Zimmerman, V., & Soliday, J. (1992). Intrapersonal and situational predictors of coaching burnout. *Journal of Sports and Exercise Psychology*, *14*, 40–58.



Pain Management

Dennis C. Turk
University of Washington, Seattle, Washington, USA

1. Pain: Magnitude of the Problem
2. Pharmacological and Surgical Treatments for Pain
3. Factors Influencing the Perception of Pain
4. Psychological Treatment of Chronic Pain Patients
5. Interdisciplinary Pain Rehabilitation Programs
6. Subgroups of Chronic Pain Sufferers
7. Concluding Comments
Further Reading

GLOSSARY

biofeedback A procedure that combines voluntary control of autonomic nervous system activity with continuous recording and presentation of information ("feedback") regarding activity of (e.g., muscles) believed to be involved in various pain syndromes; the physiological information provided to the patient by means of electrical activity converted into signals, such as tones and lights, is used by the patient to help him or her learn how to control the autonomic activity being monitored.

cognitive–behavior therapy (applied to chronic pain) Involves multiple interventions that target behavioral contributors to pain and the maladaptive thoughts that can accompany adjustment to any chronic physical condition; the focus of therapy centers around four central components: (a) education to increase the patient's knowledge about his or her condition, reassurance, and reduction of fear; (b) behavioral exercises and activities designed to increase physical and functional activities such as pacing programs to enhance physical exercise and activity; (c) cognitive exercises and activities designed to help the patient alter responses to stressors and pain; and (d) skills training to help the patient deal with problems created by a chronic disease and to modify the perception of pain using techniques such as relaxation, biofeedback, and other self-control strategies.

implantable drug delivery systems Consist of the implantation of a pump and reservoir containing analgesic medication (usually opioids); the pump provides a steady administration of the medication following a preprogrammed schedule directly into the spinal canal that may be more potent than oral administration.

interdisciplinary pain rehabilitation programs Comprehensive treatment programs that integrate physical and occupational therapy, medication management, vocational counseling, and psychological (e.g., operant conditioning, cognitive–behavior therapy) components, with an emphasis on self-management and functioning improvements despite residual pain rather than cure.

operant conditioning (applied to chronic pain) Has the main focus of modifying the frequency of a given behavior; if the consequence of the given behavior is rewarding, the likelihood of its occurrence increases, whereas if the consequence is aversive, the likelihood of its occurrence decreases.

spinal cord stimulation Involves surgical placement of electrodes along the spinal cord connected to an external device capable of generating a current along the electrodes, with the current designed to interrupt or mask the transmission of noxious sensations from the periphery to the brain.

Pain is an extremely common perceptual event that is essential for survival. However, when pain persists for extended periods of time, it no longer serves as a warning function, it affects all aspects of a person's functioning, and it can become a problem in its own

©2004 Elsevier Inc.
All rights reserved.

right. Despite advances in understanding of the anatomy and physiology of pain and the development of treatments based on sophisticated technology, chronic pain continues to be a prevalent and costly problem. Because chronic pain is a perceptual event, psychosocial and behavioral factors as well as physical ones contribute to the experience. Over time, cognitive, affective, and environmental factors play an increasingly greater role in the maintenance of pain, emotional distress, and functional disability. An understanding of the role of psychological factors has resulted in innovative treatment approaches that have been shown to have significant effects on pain intensity, associated physical symptoms, and physical functioning. These approaches are not cures; rather, they focus on functional restoration, alleviation of emotional distress, and reduction of disability. The effectiveness of psychological treatments for diverse chronic pain syndromes has been demonstrated in numerous studies. Moreover, psychological principles and techniques have been incorporated into successful interdisciplinary pain rehabilitation programs.

1. PAIN: MAGNITUDE OF THE PROBLEM

In 1998, Abbott and Fraser noted that pain is the single most frequent reason for physician consultation, with nearly one-half of those who seek treatment reporting pain as their primary symptom. Precise estimates of the prevalence of chronic pain syndromes are difficult to ascertain. In 1994, Joranson and Lietman indicated that there may be more than 30 million people with chronic or recurrent painful conditions in the United States. Based on a survey conducted by Marketdata Enterprises in 1995, an estimated 2.9 million Americans (1.1% of the population) are treated annually by health care professionals specializing in chronic pain.

The U.S. Bureau of the Census in 1996, as well as the National Research Council in 2001, estimated that the total annual costs of chronic pain (e.g., treatment, lost work days, disability payments, legal fees) in the United States range from $150 billion to $215 billion. In 1995, Cousins suggested that the costs of health care for patients with chronic pain might exceed the combined costs of treating patients with coronary artery disease, cancer, and acquired immune deficiency syndrome (AIDS).

2. PHARMACOLOGICAL AND SURGICAL TREATMENTS FOR PAIN

The first-line treatment for pain consists of a host of pharmacological agents. In 1998, Schappert noted that pain medications were the second most prescribed drugs (after cardiac–renal drugs) during visits to physicians' offices and emergency rooms, and the results of the National Ambulatory Care Medical Survey conducted in 1998 suggested that pain accounted for 12% of all medications prescribed during ambulatory office visits in the United States. Pharmaceutical industry data provided in a personal communication indicate that more than 312 million prescriptions for analgesic medications were written in the year 2000 in the United States, that is, more that one prescription for analgesics for every man, woman, and child.

The most frequently used medications to control pain in patients with chronic pain are nonsteroidal anti-inflammatory drugs, acetaminophen, opioid and nonopioid analgesics, low-dose tricyclics, and anticonvulsants. Despite their frequent use, currently available medications do not eliminate pain. For example, in 2002, Turk reviewed the available literature and concluded that the average pain reduction for patients placed on the most potent agents, opioids, is approximately 35%.

Persistence of some chronic pain syndromes (e.g., back pain) frequently leads to surgery. However, a number of studies, including those by North and colleagues in 1991 and Franklin and colleagues in 1994, reveal that significant pain may persist following surgery and that a significant percentage of chronic pain patients treated with surgery report that their pain is actually worse following surgery. Several studies of implantable devices (e.g., spinal cord stimulators, implantable drug administration systems) have reported impressive results in pain reduction, but only for carefully selected patients with pain of long duration.

3. FACTORS INFLUENCING THE PERCEPTION OF PAIN

There has been a growing recognition that pain is a complex perceptual experience influenced by a wide range of psychosocial factors, including emotions, social and environmental contexts, sociocultural background, the meaning of pain to the person, and beliefs, attitudes, and expectations, as well as by

biological factors. Biological factors may initiate, maintain, and modulate physical perturbations, but psychological factors influence the appraisal and perception of internal physiological signs and social factors shape the behavioral responses of patients to the perceptions of their physical perturbations and may have a direct effect on physiological processes associated with exacerbation and maintenance of symptoms.

Pain that persists for months or years, called chronic pain, will have an impact on all aspects of a person's functioning—emotional, interpersonal, avocational, and physical. Consequently, treating chronic pain patients successfully requires attention not only to the organic basis of the symptoms but also to the range of factors that modulate nociception and moderate the pain experience and related disability.

4. PSYCHOLOGICAL TREATMENT OF CHRONIC PAIN PATIENTS

A large volume of research exists demonstrating the efficacy of the psychological treatments for low back pain, fibromyalgia syndrome, noncardiac chest pain, arthritis, headaches, temporomandibular disorder (TMD), and whiplash-associated disorders, to name a few. Furthermore, cognitive–behavioral approaches appear to prevent the development of disability due to pain. Psychological modalities have been used in the treatment of chronic pain patients either on their own or, more typically, as an essential component of comprehensive interdisciplinary rehabilitation.

Table I lists the most common psychological approaches applied to chronic pain. The effectiveness of relaxation, biofeedback, and cognitive–behavior therapy (CBT) has been the most thoroughly evaluated. Electromyographic (EMG) biofeedback is a procedure that combines muscle relaxation with continuous recording and presentation of information (i.e., "feedback") regarding activity of muscles believed to be involved in different pain syndromes. In 1979, Turk and colleagues noted that biofeedback is based on the idea that stress-induced hyperactivity is often an important component in muscular pain syndromes. Biofeedback makes use of surface records of muscle activity to provide patients with information about levels of muscle tension. Levels of muscular arousal are recorded and converted into signals (usually auditory or visual) that permit patients to observe their current levels of muscle tension. The signals provided indicate changes in muscular activity resulting from patients' efforts to relax the muscles. Thus, the information about muscular activity is "fed back" to the patients and assists them in learning to alter high levels of muscular tension that may be contributing to the onset, maintenance, or exacerbation of pain.

TABLE I
Treatments for Chronic Pain

Pharmacological
Nonsteroidal anti-inflammatory agents
Acetaminophen
Opioids
Antidepressants
Anticonvulsants (e.g., gabapentin)
Topical agents (e.g., lidocaine)
Physical modalities (e.g., hot packs, ultrasound, exercise, transcutaneous nerve stimulation)
Surgery
Regional anesthesia (e.g., neural blockade, epidural steroid injections)
Spinal cord stimulators
Psychological
Relaxation
Electromyographic biofeedback
Hypnosis
Counseling (e.g., family, marital)
Behavior modification (e.g., exposure, desensitization, contingency contracting, operant conditioning)
Cognitive–behavior therapy
Complementary and alternative medicine (e.g., acupuncture, chiropracty, diet supplements, Tai chi)

The main focus of operant conditioning is modification of the frequency of a given behavior. If the consequence of the given behavior is rewarding, the likelihood of its occurrence increases; if the consequence is aversive, the likelihood of its occurrence decreases. Behaviors associated with pain, such as limping and moaning, are called "pain behaviors"—overt expressions of pain, distress, and suffering. When an individual is exposed to a stimulus that causes tissue damage, the immediate behavior is withdrawal in an attempt to escape from noxious sensations. Such pain behaviors are adaptive and appropriate. Over time, however, environmental contingencies of reinforcement may serve to maintain the pain behaviors long after the initial cause of the pain has been resolved (Table II). The behaviors are

TABLE II
Operant Conditioning: Modification of Behavior

Schedule	*Consequences*	*Probability of the behavior recurring*
Positive reinforcement	Reward the behavior	More likely
Negative reinforcement	Prevent or withdraw aversive results	More likely
Punishment	Punish the behavior	Less likely
Neglect	Prevent or withdraw positive results	Less likely

conditioned based on their consequences. Often, the consequences are provided by significant others, including health care providers, who may inadvertently reinforce the pain behaviors by providing attention, treatment, and/or avoidance of activities that are undesirable to the patient. The operant learning paradigm does not uncover the etiology of pain but rather focuses primarily on the maintenance of pain behaviors and on deficiency in well behaviors such as activity. Treatment focuses on extinction of pain behaviors and on increased well behaviors by changing the reinforcement contingencies—withdrawal of rewards for pain behaviors and increased rewards for well behaviors.

4.1. Cognitive–Behavior Therapy: Components

CBT programs usually involve multiple interventions that target behavioral contributors to pain and the maladaptive thoughts that can accompany adjustment to any chronic physical condition. In 1999, Turk and Okifuji found that these programs focus on four primary components: (a) education to increase patients' knowledge about their conditions, reassurance, and reduction of fear; (b) behavioral exercises and activities designed to increase physical and functional activities such as pacing programs to enhance physical exercise and activity; (c) cognitive exercises and activities designed to help patients alter their responses to stressors and pain; and (d) skills training to help patients deal with problems created by chronic diseases and to modify the perception of pain using techniques such as relaxation, biofeedback, and other self-control strategies (Table III).

TABLE III
Components of Cognitive–Behavior Therapy

- Provide education about pain and particular syndrome.
- Encourage acceptance.
- Foster a self-management perspective.
- Focus on function rather than cure.
- Assist in goal setting.
- Teach relaxation.
- Provide information and guidance about pacing and increasing activities.
- Provide guidance on ways in which to improve sleep.
- Emphasize identifying and eliminating maladaptive thoughts.
- Provide strategies to help cope with relapse.
- Provide skills and skills training for ways in which to cope with interpersonal problems.
- Discuss appropriate uses of medications, exercise, and physical modalities (e.g., heat, cold).

4.2. Cognitive–Behavior Therapy: Efficacy

In 1999, Morley and colleagues conducted a systematic review and meta-analysis of randomized controlled trials, comparing the effectiveness of CBT to waiting list controls and alternative treatment control conditions. They found that CBT produced significantly greater changes in pain experience, cognitive coping, and reduced behavioral expressions of pain. However, significant differences were not found for dysphoric mood, negative appraisals, and social role functioning.

5. INTERDISCIPLINARY PAIN REHABILITATION PROGRAMS

The large body of evidence demonstrating the importance of psychosocial and behavioral factors has led to the proliferation of interdisciplinary pain rehabilitation programs (IPRPs). In 2001, Loeser and Turk noted that these IPRPs typically include psychological (e.g., operant conditioning, CBT) components as integrated within comprehensive treatment plans that include physical and occupational therapy, medication management, and vocational counseling.

In 1992, Flor and colleagues concluded that IPRPs are more effective than no treatment, waiting list controls, and monodisciplinary treatments in terms of reductions of health care use and medication,

increased activity and return to work, closure of disability claims, and reductions of affective distress. The failure to find significant improvement in affective distress, noted by Morley and colleagues in 1999, is at variance with the results of this meta-analysis. The explanation for this inconsistency is unclear and may involve differences in the effects of psychological treatments on their own compared with when they are integrated within IPRPs. IPRPs are often referred to as being based on operant, behavioral, or cognitive–behavioral principles. The actual treatments are broadly defined and often include diverse ingredients. It is difficult to determine what the active components of the treatment packages actually are. Turk in 1990, as well as Kole-Snijders and colleagues in 1999, indicated that this is important because it is not cost-effective to provide comprehensive treatments with multiple components, some of which may be unnecessary. Patients with different characteristics may benefit from different treatments, and not all of the components included may be necessary for all patients.

The reduction of pain following treatment at IPRPs has been reported to be statistically significant in meta-analyses conducted by Flor and colleagues in 1992. Flor and colleagues reported the mean pain reduction for patients treated at IPRPs at 37%. However, the majority of patients continue to experience considerable pain. Interestingly, Flor and colleagues found that the pain reduction achieved at IPRPs, which is comparable to the pain reduction reported with opioids, is accompanied with a significant decrease (63%) in prescription pain medication. However, even when successful, people continue to experience significant levels of pain with which they must cope.

6. SUBGROUPS OF CHRONIC PAIN SUFFERERS

Although biobehavioral treatments appear to be important components of rehabilitation treatments for chronic pain patients, not all patients benefit equally. Focusing solely on group effects may mask important issues related to the characteristics of patients who respond successfully to a treatment. Chronic pain syndromes are made up of heterogeneous groups of people even if the individuals have the identical medical diagnosis. A common pitfall in clinical research and practice is the assumption of patient homogeneity. Inclusion of a diverse group of patients into the same category simply because they present a common set of symptoms may result in inconsistent research results, as is commonly observed in the pain literature.

Several studies, including one by Kight and colleagues in 1999 and two by Turk and colleagues in 1998, have identified subgroups of patients based on psychosocial and behavioral characteristics. In 1997, Dahlstrom and colleagues found that when patients were classified into different subgroups based on their psychosocial and behavioral responses, they responded differentially to treatments. Similarly, Turk and colleagues noted differential responses to a common treatment for patients with distinctive psychological characteristics but identical physical diagnoses. The results of these studies could be used for developing treatments that are matched to patients' characteristics. Various studies comparing the subgroups have yielded evidence supporting differential responses to the same intervention.

7. CONCLUDING COMMENTS

Despite the advances in the knowledge of sensory physiology, anatomy, and biochemistry and of the prevalence of chronic pain, and despite the many treatments available to provide relief (see Table I), total elimination of pain is elusive for many pain sufferers. Pain and disability have not been eliminated as problems for a significant portion of the population and for society as a whole. Comprehensive biobehavioral treatments appear to have the best outcomes for patients with many chronic pain syndromes. Combining CBT with more traditional medical, surgical, and pharmacological modalities appears to be the most reasonable approach. Matching treatments with patient characteristics has the potential to improve outcomes and reduce costs.

Acknowledgments

Preparation of this article was supported in part from grants provided by the National Institute of Arthritis and Musculoskeletal and Skin Diseases (AR/AI44724 and AR47298), the National Institute of Child Health and Human Development/National Center for Medical Rehabilitation Research (HD33989), and the Fetzer Institute.

See Also the Following Articles

Coping ■ Psychological Rehabilitation Therapies

Further Reading

Abbott, F. V., & Fraser, M. I. (1998). Use and abuse of over-the-counter analgesic agents. *Journal of Psychiatry and Neuroscience*, *23*, 13–34.

Cousins, M. J. (1995). Foreword. In W. E. Fordyce (Ed.), *Back pain in the workplace: Management of disability in nonspecific conditions* (p. ix). Seattle, WA: IASP Press.

Dahlstrom, L., Widmark, G., & Carlsson, S. G. (1997). Cognitive–behavioral profiles among different categories of orofacial pain patients: Diagnostic and treatment implications. *European Journal of Oral Science*, *105*, 377–383.

Dworkin, S. F. (1995). Behavioral characteristics of chronic temporomandibular disorders: Diagnosis and assessment. In B. J. Sessle, P. S. Bryant, & R. A. Dionne (Eds.), *Progress in pain research and management* (pp. 175–192). Seattle, WA: IASP Press.

Dworkin, S. F., Turner, J. A., Wilson, L., Massoth, D., Whitney, C., Huggins, K. H., Burgess, J., Sommers, E., & Truelove, E. (1994). Brief group cognitive–behavioral intervention for temporomandibular disorders. *Pain*, *59*, 175–187.

Epker, J., & Gatchel, R. J. (2000). Coping profile differences in the biopsychosocial functioning of patients with temporomandibular disorders. *Psychosomatic Medicine*, *62*, 69–75.

Flor, H., Fydrich, T., & Turk, D. C. (1992). Efficacy of multidisciplinary pain treatment centers: A meta-analytic review. *Pain*, *49*, 221–230.

Franklin, G. M., Haug, J., & Heyer, N. J. (1994). Outcome of lumbar fusion in Washington State Workers' Compensation. *Spine*, *19*, 1897–1904.

Gatchel, R. J., & Turk, D. C. (1999). *Psychosocial factors in pain: Critical perspectives*. New York: Guilford.

Goossens, M., Rutten-Van Molken, M., Kole-Snijders, A., Vlaeyen, J., van Breukelen, G., & Leidl, R. (1998). Health economic assessment of behavioral rehabilitation in chronic low back pain: A randomized clinical trial. *Health Economics*, *7*, 39–51.

Hasenbring, M., Ulrich, H. W., Hartmann, M., & Soyka, D. (1999). The efficacy of a risk factor-based cognitive behavioral intervention and electromyographic biofeedback in patients with acute sciatic pain: An attempt to prevent chronicity. *Spine*, *24*, 2525–2535.

Hildebrandt, J., Pfingsten, M., Saur, P., & Jansen, J. (1997). Prediction of success from a multidisciplinary program for chronic low back pain. *Spine*, *22*, 990–1001.

Holroyd, K., & Lipchik, G. (1999). Psychological management of recurrent headache disorders: Progress and prospects. In R. J. Gatchel, & D. C. Turk (Eds.), *Psychosocial factors in pain: Critical perspectives* (pp. 193–212). New York: Guilford.

Joranson, D. E., & Lietman, R. (1994). *The McNeil National Pain Survey*. New York: Louis Harris & Associates.

Kight, M., Gatchel, R. J., & Wesley, L. (1999). Temporomandibular disorders: Evidence for significant overlap with psychopathology. *Health Psychology*, *18*, 177–182.

Kole-Snijders, A. M. J., Vlaeyen, J. W. S., Goossens, M. E. J. B., Rutrten-van Molken, M. P. M. H., Heuts, P. H. T. G., van Breukelen, G., & van Eck, H. (1999). Chronic low-back pain: What does cognitive coping skills training add to operant behavioral treatment? Results of a randomized clinical trial. *Journal of Consulting and Clinical Psychology*, *67*, 931–944.

Loeser, J. D., & Turk, D. C. (2001). Multidisciplinary pain management. In J. D. Loeser, S. D. Butler, C. R. Chapman, & D. C. Turk (Eds.), *Bonica's management of pain* (3rd ed., pp. 2069–2087). Philadelphia: Lippincott, Williams, & Wilkins.

Lorig, K., Mazonson, P., & Holman, H. (1993). Evidence suggesting that health education for self-management in patients with chronic arthritis has sustained health benefits while reducing health care costs. *Arthritis & Rheumatism*, *36*, 439–446.

Main, C. J., & Spanswick, C. C. (2000). *Pain management: An interdisciplinary approach*. London: Churchill–Livingstone.

Marketdata Enterprises. (1995). Chronic pain management programs: A market analysis. Valley Stream, NY: Author.

Mayou, R. A., Bryant, M. M., Sanders, D., Bass, C., Klimes, I., & Farfar, C. (1997). A controlled trial of cognitive–behavioural therapy for non-cardiac chest pain. *Psychological Medicine*, *27*, 1021–1032.

Morley, S., Eccleston, C., & Williams, A. (1999). Systematic review and meta-analysis of randomized controlled trials of cognitive behaviour therapy and behaviour therapy for chronic pain in adults, excluding headache. *Pain*, *80*, 1–13.

National Center for Health Statistics. (1998). *National Ambulatory Medical Care Survey, 1998*. Washington, DC: Department of. Health and Human Services.

National Research Council. (2001). *Musculoskeletal disorders and the workplace*. Washington, DC: National Academy Press.

North, R. G., Ewend, M. G., & Lawton, M. T. (1991). Failed back surgery syndrome: Five-year follow-up after spinal cord stimulator implantation. *Neurosurgery*, *28*, 692–699.

Schappert, S. M. (1998). *Ambulatory care visits to physicians' offices, hospital outpatient departments, and emergency departments: United States, 1996* (Vital and Health Statistics, Series 13, No. 134). Atlanta, GA: Centers for Disease Control and Prevention.

Turk, D. C. (1990). Customizing treatment for chronic pain patients: Who, what, and why. *Clinical Journal of Pain*, *6*, 255–270.

Turk, D. C. (2002). Clinical effectiveness and cost effectiveness of treatments for chronic pain patients. *Clinical Journal of Pain*, *18*, 355–365.

Turk, D. C., & Gatchel, R. J. (2002). *Psychological approaches to pain management: A practitioner's handbook* (2nd ed.). New York: Guilford.

Turk, D. C., Meichenbaum, D., & Berman, W. H. (1979). The application of biofeedback for the regulation of pain: A critical review. *Psychological Bulletin*, *86*, 1322–1338.

Turk, D. C., & Melzack, R. (2001). *Handbook of pain assessment* (2nd ed.). New York: Guilford.

Turk, D. C., & Okifuji, A. (1999). A cognitive–behavioral approach to pain management. In P. D. Wall, & R. Melzack (Eds.), *Textbook of pain* (4th ed., pp. 1431–1444). London: Churchill–Livingstone.

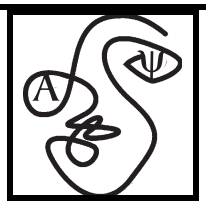
Turk, D. C., & Okifuji, A. (2002). Psychological factors in chronic pain: Evolution and revolution. *Journal of Consulting and Clinical Psychology*, *70*, 278–290.

Turk, D. C., Okifuji, A., Sinclair, J. D., & Starz, T. W. (1998). Interdisciplinary treatment for fibromyalgia syndrome: Clinical and statistical significance. *Arthritis Care and Research*, *11*, 186–195.

Turk, D. C., Okifuji, A., Starz, T. W., & Sinclair, J. D. (1998). Differential responses by psychosocial subgroups of fibromyalgia syndrome patients to an interdisciplinary treatment. *Arthritis Care and Research*, *11*, 397–404.

U.S. Bureau of the Census. (1996). *Statistical abstract of the United States: 1996* (116th ed.). Washington, DC: Author.

Vendrig, A. A., Van Akkerveeken, P. F., & McWhorter, K. R. (2000). Results of multimodal treatment program for patients with chronic symptoms after whiplash injury of the neck. *Spine*, *25*, 238–244.



Panic

Martin M. Antony

St. Joseph's Healthcare and McMaster University, Hamilton, Ontario, Canada

1. Introduction
2. Definitions and Description
3. Epidemiology of Panic Disorder
4. Causes of Panic Disorder
5. Treatment of Panic Disorder
Further Reading

Panic is an emotional state associated with extreme fear and discomfort, a sense of dread or apprehension, and a number of intense physical sensations, such as racing heart, sweating, shaking, shortness of breath, and various other symptoms of arousal. It tends to begin suddenly, is often triggered by a feared situation or object, and may occur out of the blue, without any obvious trigger or cause.

GLOSSARY

agoraphobia A fear of situations in which escape might be difficult, or in which help might be unavailable, in the event of experiencing a panic attack or panic-like sensations.

anxiety sensitivity A fear of anxiety-related symptoms, such as dizziness, shortness of breath, and racing heart.

cognitive behavior therapy A form of psychotherapy used to change negative patterns of thinking and maladaptive behaviors using cognitive restructuring, exposure-based interventions, relaxation training, and other methods.

interoceptive exposure Exposing oneself to feared physical symptoms using exercises designed to trigger the sensations.

panic attack A rush of fear or discomfort that peaks quickly and that is associated with intense symptoms of physical arousal, including racing heart, shaking, and feeling lightheaded.

panic disorder An anxiety disorder in which an individual experiences recurrent unexpected panic attacks, accompanied by anxiety over having more attacks, worry about the consequences of the attacks, and changes in behavior aimed at preventing future attacks.

1. INTRODUCTION

The word panic has its origins in Greek mythology. It refers to Pan, the ancient Greek god of woods, fields, and flocks, who has a human torso and head, and the legs, horns, and ears of a goat. Pan was believed to be very mischievous, often inciting terror among travelers making their way through the forest by rustling trees and bushes. Pan's victims probably experienced many of the same sensations that people report during modern day panic attacks.

The experience of panic is often referred to as a "fight or flight" response, in which the body is mobilized to facilitate immediate escape from a perceived threat, or to meet the threat head-on with an aggressive response. During panic, the heart races to get blood to the large muscles that may be needed in the event of an escape. Breathing becomes more rapid in order to increase the flow of oxygen to the bodily tissues. Perspiration increases to cool off the body so it can perform more efficiently. In fact, almost all of the

©2004 Elsevier Inc.
All rights reserved.

symptoms that occur during a typical episode of panic can be explained in terms of their adaptive functions. In other words, the symptoms of panic serve to improve an individual's ability to escape from perceived danger.

However, panic is often not adaptive or helpful. In many cases, panic occurs in the absence of true danger or threat. Further, panic attacks may lead individuals to avoid feared situations, and may therefore cause considerable impairment in functioning. As a result, frequent panic attacks often lead an individual to seek treatment in one form or another.

2. DEFINITIONS AND DESCRIPTION

In the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV), a panic attack is described as a discrete period of intense fear or discomfort that peaks in 10 minutes or less, and is accompanied by at least four of the following 13 symptoms: (1) racing or pounding heart, (2) sweating, (3) trembling or shaking, (4) shortness of breath, (5) feeling of choking, (6) chest pain or discomfort, (7) nausea or abdominal distress, (8) feeling dizzy, unsteady, or faint, (9) feeling unreal or detached, (10) paresthesias (i.e., numbness or tingling sensations), (11), chills or hot flushes, (12) fear of dying, and (13) fear of going crazy or losing control.

Panic attacks are usually triggered by a feared object, situation, or thought. For example, a person who fears dogs may experience a panic attack each time he or she encounters a dog, or panic attacks may occur in anticipation of giving a speech for someone who is frightened of public speaking. Similarly, a person with obsessive compulsive disorder may experience panic attacks in response to obsessional thoughts (e.g., irrational thoughts about harming another individual).

Although panic attacks are often triggered, they can also appear unexpectedly or out of the blue. The presence of recurrent, unexpected panic attacks is the hallmark feature of panic disorder. In addition, people with panic disorder tend to be fearful of experiencing panic attacks and the symptoms that are associated with their attacks. They often anticipate the attacks, wondering when the next one will occur. They also are concerned about the possible consequences of the attacks, worrying about a physical catastrophe such as a heart attack, a stroke, dying, or fainting, or about an emotional or psychological catastrophe such as losing control, embarrassing oneself, or "going crazy." Finally, people with panic disorder often change their behavior in response to the attacks. They may actively avoid situations where the attacks occur, and they may do things to protect themselves from experiencing further attacks (e.g., distraction, avoiding being alone, checking their body for panic symptoms, carrying medication). To receive a diagnosis of panic disorder, it must also be established that the attacks are not due to a medical illness (e.g., cardiac disease or hyperthyroidism) or a substance (e.g., cocaine use, caffeine overuse, alcohol withdrawal).

Most people with panic disorder also develop some degree of agoraphobia. Agoraphobia is a fear of situations in which escape might be difficult, or in which help might be unavailable, in the event of experiencing a panic attack or panic-like sensations. Typical situations avoided by people with agoraphobia include driving, flying, crowds, enclosed places, public transportation, malls, supermarkets, and traveling. In extreme cases, people with agoraphobia may avoid leaving the house altogether. Often, agoraphobic fear is milder when the individual is accompanied by a close friend, relative, or spouse.

3. EPIDEMIOLOGY OF PANIC DISORDER

Studies on the prevalence of panic disorder estimate that between 2 and 5% of individuals in the general population suffer from this problem, although up to one-third of individuals may experience a panic attack from time to time. Panic disorder occurs more frequently in women than in men. In addition, women with panic disorder tend to experience more severe symptoms (e.g., more intense and frequent attacks, more severe agoraphobic avoidance) than men with panic disorder. This condition most often begins in early adulthood (in a person's 20s or early 30s), though it can begin at any age, ranging from childhood to late in life. The onset often occurs following a period of stressful life events. Panic disorder occurs across ethnic and religious groups and does not appear to be significantly associated with intelligence, marital status, education, or socioeconomic status.

Panic disorder is often accompanied by significant functional impairment and leads to considerable costs to society in the form of increased healthcare utilization, inability to work, and decreased productivity at work. Without treatment, the course is usually

chronic, with up to 90% of individuals still experiencing panic attacks a year after their initial assessment. Panic disorder often occurs concurrently with other psychological problems, including other anxiety disorders and depression. The rate of substance use disorders is also higher in people with panic disorder than it is in the general population.

4. CAUSES OF PANIC DISORDER

There is evidence that both psychological and biological factors contribute to the development and maintenance of panic disorder. From a psychological perspective, cognitive factors seem to play an important role. A risk factor for the development of panic disorder is heightened anxiety sensitivity (i.e., a fear of experiencing symptoms of anxiety and arousal). In addition, people with panic disorder tend to misinterpret the meaning of their physical symptoms, and often assume that their symptoms are dangerous when in fact they are not. Panic disorder is also associated with a tendency to attend to threat-related information (e.g., to scan the body for uncomfortable symptoms) and to remember information that is consistent with anxiety-related beliefs and assumptions. It is believed that the tendency to pay attention to threat-related cues and to interpret benign physical symptoms as dangerous contributes to the development of panic disorder and to the chronic course of this condition.

From a biological perspective, a number of different factors have been studied and found to be relevant. Neurotransmitters in the brain (e.g., norepinephrine, serotonin, cholecystokinin) may play a role. For example, substances that increase levels of norepinephrine or cholecystokinin in the brain (including injections of cholecystokinin and inhalation of carbon dioxide-enriched air) have been found to trigger panic attacks, and medications that block the re-uptake of norepinephrine and serotonin are effective for preventing panic attacks. Genetics also influences the development of this condition. Panic disorder tends to run in families, and a portion of the transmission of panic disorder from generation to generation is directly attributable to the effects of genetics. Finally, brain imaging studies suggest that certain areas in the brain (e.g., the hippocampus) may be more or less active depending on whether someone has panic disorder. The nature of these differences is still not fully understood, and research findings are often inconsistent.

5. TREATMENT OF PANIC DISORDER

Evidence-based approaches to treating panic disorder include cognitive-behavior therapy (CBT), medications, or a combination of these approaches. On average, these options are about equally effective in the short term. However, any one individual may benefit more from one approach over another. In the long term, there is evidence favoring CBT. Rates of relapse and recurrence tend to be lower following the discontinuation of CBT than they are following the discontinuation of medication. Although most people benefit considerably from these treatments, fewer than half of individuals experience a nearly complete remission of their symptoms. Occasional panic attacks, anxiety, and avoidance are common following the end of treatment.

CBT is designed to help an individual to change the beliefs and behaviors that are thought to maintain the problem. The client is taught to replace anxious and unrealistic thoughts with more realistic beliefs, by critically examining the evidence for both anxious and non-anxious interpretations. Clients are also taught to confront feared situations by engaging in prolonged exposure on a regular basis (e.g., practicing driving for an hour each day) until the situations no longer trigger fear. They also learn to confront the fear of physical sensations through interoceptive exposure, that is, exposure to arousal symptoms (e.g., spinning to induce dizziness, aerobic exercise to increase one's heart rate). Finally, relaxation-based strategies (especially slow diaphragmatic breathing) are sometimes used to reduce the intensity of physical symptoms, although it is not clear that these strategies provide any benefit over and above the effects of the other CBT techniques. Usually, education about the nature and treatment of panic and anxiety is included throughout the treatment.

The most commonly used medications for panic disorder include certain antidepressants as well as a number of anxiolytics. Effective antidepressants include the selective serotonin re-uptake inhibitors (e.g., citalopram, fluoxetine, paroxetine, sertraline), certain tricyclic antidepressants (e.g., imipramine, clomipramine), as well as other medications (e.g., venlafaxine). Anxiolytics, such as the benzodiazepines (e.g., alprazolam, clonazepam) have also been shown to reduce the symptoms of panic disorder. Decisions regarding which medication to use should be based on the individual's past treatment history, medical history, possible interactions between the medication and other drugs

that person may be taking, potential side effects, and other factors.

See Also the Following Articles

Anxiety Disorders in Late Life ■ Test Anxiety

Further Reading

Antony, M. M. (2001). Measures for panic disorder and agoraphobia. In M. M. Antony, S. M. Orsillo, & L. Roemer (Eds.), *Practitioner's guide to empirically-based measures of anxiety* (pp. 95–126). New York, NY: Kluwer Academic/Plenum.

Antony, M. M., & Swinson, R. P. (2000). *Phobic disorders and panic in adults: A guide to assessment and treatment*. Washington, DC: American Psychological Association.

Craske, M. G., & Barlow, D. H. (2000). *Mastery of your anxiety and panic,* 3rd ed. Boulder, CO: Graywind Publications, Inc.

Craske, M. G., & Barlow, D. H. (2001). Panic disorder and agoraphobia. In D. H. Barlow (Ed.), *Clinical handbook of psychological disorders,* 3rd ed. New York: Guilford Press.

McNally, R. J. (1994). *Panic disorder: A critical analysis*. New York: Guilford Press.

Pollard, C. A., & Zuercher-White, E. (2003). *The agoraphobia workbook: A comprehensive program to end your fear of symptom attacks*. Oakland, CA: New Harbinger.

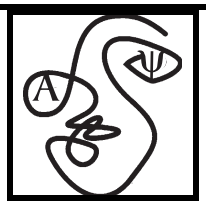
Rachman, S., & de Silva, P. (1996). *Panic disorder: The facts*. New York, NY: Oxford University Press.

Rosenbaum, J. F., & Pollack, M. H. (1998). *Panic disorder and its treatment*. New York, NY: Marcel Dekker.

Taylor, S. (2000). *Understanding and treating panic disorder: Cognitive and behavioral approaches*. Chichester, UK: Wiley.

Zuercher-White, E. (1997). *An end to panic: Breakthrough techniques for overcoming panic disorder*, 2nd ed. Oakland, CA: New Harbinger Publications.

Zuercher-White, E. (1997). *Treating panic disorder and agoraphobia: A step by step clinical Guide*. Oakland, CA: New Harbinger Publications.



Paralinguistic Behaviors and Culture

Hede Helfrich

Universität Hildesheim, Hildesheim, Germany

1. Introduction
2. The Nature of Paralinguistic Behaviors
3. The Relevance of Culture
4. Evolutionary Continuity in Communication
5. Paradigms for the Study of Paralinguistic Behaviors
6. Externalization of Paralinguistic Behaviors
7. Accuracy of Listeners' Inferences
8. Use of Paralinguistic Cues by Listeners
9. Conclusion

Further Reading

GLOSSARY

baby talk A voice and speech style used by adults, and even by children, when speaking to infants.

contact cultures Cultures from which people stand closer, face more directly, touch, and look more than do people from noncontact cultures.

distance between bodies Physical space between people in their conversations; there are wide cultural differences in what is considered appropriate distance.

facial expression Patterns of movements of brow, eye, and mouth regions.

paralinguistic behaviors Vocal and nonvocal phenomena that are either blended or associated with speech utterances but have no direct function according to linguistic rules; they include tone of voice, speech and conversation style, facial expressions, gestures, head and body movements, and distance between bodies.

silence The absence of vocal behavior for a period of time; often, it is misunderstood to refer to the absence of any and all communication.

tone of voice Pitch, loudness, and timbre of a spoken utterance.

turn taking The allocation of speaker and listener roles in a conversation; a person takes his or her turn when adopting the role of the speaker.

Successful communication includes knowing when to speak and when to remain silent; which tone of voice, which facial expression, and which body position are appropriate in various contexts; what the routines for turn taking are in conversation; and how to exhibit the appropriate listener signals. These behaviors are referred to as "paralinguistic." Some of them may be common to all people, and some of them may vary from one culture to another. When communicating with persons who are culturally different, violations of the rules for paralinguistic behaviors often lead to misunderstandings, even if the words and grammar of the foreign language are appropriately mastered. The aim of the cross-cultural study of paralinguistic behaviors is to learn how the principles of paralinguistic behaviors vary from one culture to another and how universal principles must be qualified and adjusted on the basis of cultural differences. This understanding will help people to use communication effectively in intercultural encounters.

©2004 Elsevier Inc.
All rights reserved.

1. INTRODUCTION

One of the most striking differences among members of various cultures is that they speak different languages. Thus, it will obviously be useful to persons involved in international transactions to learn the languages of their interlocutors. However, successful communication involves knowing not only what to say to whom but also how to say it appropriately in any given situation. It includes knowing when to speak and when to remain silent; which tone of voice, which facial expression, and which body position are appropriate in various contexts; what the routines for turn taking are in conversation; how to request, how to offer, and how to decline assistance or cooperation; and how to exhibit the appropriate listener signals. These behaviors are not regarded by linguists as being an integral part of the verbal utterances with which they are associated; therefore, they are referred to as "paralinguistic." Some of these behaviors may be universal (common to all people), and some of them may vary from one culture to another. As a person learns his or her "mother tongue" during childhood, that person usually acquires and reproduces the appropriate paralinguistic behaviors typical of speakers of that language because a first language is learned in the context of behavioral settings where the norms of the culture are acted and enunciated. However, these norms are less explicit and less under conscious control than are the rules of languages. Thus, learning a second language has traditionally stressed the acquisition of words and grammar while ignoring paralinguistic behaviors. When communicating with persons who are culturally different, violations of the rules for paralinguistic behaviors often lead to misunderstandings, even if the words and grammar of the foreign language are appropriately mastered. The aim of the cross-cultural study of paralinguistic behaviors is to learn how the principles of paralinguistic behaviors vary from one culture to another and how universal principles must be qualified and adjusted on the basis of cultural differences. This understanding will help people to use communication effectively in intercultural encounters.

2. THE NATURE OF PARALINGUISTIC BEHAVIORS

2.1. Description of Phenomena

Speech utterances not only convey information via spoken words and grammatical structure but also include phenomena that, according to linguistic rules, have no direct function. These phenomena refer to the voice in which words are uttered and to their patterning in time. In face-to-face communication, speech utterances are accompanied by various body movements such as face and head movements, gestures, body postures, and body orientation. These vocal and nonvocal phenomena, which are either blended or associated with speech utterances, are called "paralinguistic behaviors." Not all vocal modifications of spoken words are nonverbal; linguists distinguish between prosodic and paralinguistic features. Prosodic features, such as stress, tone, and juncture, follow the rules of language and affect the linguistic meaning of utterances. Paralinguistic features convey other kinds of information. They not only refer to the speech behavior of a single speaker but also comprise nonverbal aspects of a conversation between two or more speakers such as regulation of speaker roles and distance between bodies. Sometimes, the term "paralinguistic" is restricted to vocal features, excluding the nonvocal characteristics that are referred to as nonverbal behaviors or body language. This article uses the broader sense, comprising both vocal and nonvocal aspects of speech, although the focus is primarily on vocal aspects.

2.2. Measurement of Paralinguistic Behaviors

2.2.1. Levels of Analysis

Paralinguistic behaviors have been analyzed at four levels: at the motor–physiological level of the body organs that produce the signals, at the physical level using acoustic or spatial parameters, at the perceptual level of the various sensory (mainly auditory and visual) modalities, and at the functional level referring to "meaningful" voice or movement categories. Analysis techniques at the first three levels may be summarized as "microanalytic techniques," whereas functional techniques may be referred to as "macroanalytic techniques." Studies using direct recordings of muscle activities such as electromyography (EMG) are rare; most of the analyses carried out are based on audio or video recordings.

In many studies, human observers are used to assess paralinguistic behaviors. For some types of studies, the human assessment is indispensable because such studies focus on the perceptual properties of the utterances or on the inferences drawn by listeners. In other studies, the judgment of human observers is used as a

more readily available or more holistic measure as opposed to the rather analytically defined physical measures. The disadvantage might be that the human system is rarely able to isolate the various physical parameters from each other. The advantage might be that sometimes information is conveyed based on complex physical configurations that are more easily judged as holistic impressions than formalized in terms of physical parameters.

2.2.2. Vocal Behavior

Human vocal behavior is the joint product of the action of three muscle systems: the respiratory, phonatory, and articulatory systems. The action of these muscle systems leads to air pressure changes, which can be perceived by listeners and described in terms of acoustic parameters based on audio recordings.

Three major acoustic dimensions are distinguished: time, frequency, and amplitude. Time-based parameters are related to the number and length of sound and silence segments within the speech flow, to onset and offset parameters of speech passages, and to speaker switching. Amplitude-based parameters refer to the energy level of voice. Of the parameters based on frequency, the one used most often is fundamental frequency (F_0). It is defined as the number of repetitions of a complex wave per second, measured in hertz (Hz), and corresponds to the frequency at which the vocal cords open and close. Averaged over a speech utterance, it can be expressed as mean F_0. Changes in F_0 may be measured in terms of F_0 variability or in terms of F_0 contour, referring to the course of consecutive F_0 values. Some important measures involve more than one dimension. For example, the voice spectrum refers to the frequency–amplitude pattern of a speech signal, showing the component frequencies of a complex waveform and their relative amplitudes.

At the perceptual level, the tone of voice, speech style, and conversation style are distinguished. Although there is no one-to-one relationship, the voice characteristics of pitch, loudness, and timbre are the perceptual counterparts of F_0, energy level, and voice spectrum, respectively. Speech style can be assessed in terms of temporal properties, such as pauses and speech tempo, and in terms of speech disfluencies, such as "filled pauses" (i.e., interruptions of continuity filled by nonlinguistic vocalizations such as "ah," "mhm," and "er") and slips of the tongue. The conversation style refers to features such as opening and ending a dialogue, speaker switching and interruptions of the speaker by another person, turn-taking, and feedback signals.

At the functional level, voices may be characterized by holistic labels such as harsh, powerful, and pleasant, and speech styles may lead to impressions such as halting and hectic speaking. Feedback signals and interruptions in conversation may signal interest or attentiveness as well as criticism or disagreement.

Table I illustrates the common measures at the acoustic and perceptual levels and gives some examples of functional characterizations of voice, speech, and conversation style.

2.2.3. Nonvocal Behavior

Analysis of nonvocal paralinguistic behaviors sometimes relies on impressionistic descriptions of real-time behavior, but it is based mostly on video recordings. The most common microanalytic technique for the analysis of facial movements is the Facial Action Coding System (FACS) developed by Ekman and Friesen in 1978. It is an anatomically based system for identifying the muscular contractions (so-called action units) in the brow, eye, and mouth regions responsible for changes in facial movement that roughly correspond to functional descriptions such as various kinds of smiling and various emotions. The original manual coding system has been automated in terms of digital feature point tracking. A microanalytic technique for the study of gestures and body movement is provided by the Bernese System, which allows a quantitative description of body and gestural positions and movements. Functional characterizations of gestures distinguish emblems, illustrators, and adaptors. Emblems directly translate words or phrases, for example, the signs for "yes" and "no." Illustrators accompany verbal utterances and clarify their verbal meanings, for example, upward movements that accompany the verbal expression "It's up there." Adaptors refer to movements that have no communicative function but serve some personal need, for example, scratching one's head until the itch is eliminated.

2.3. Possible Functions of Paralinguistic Behaviors

On the one hand, paralinguistic behavior is a product of the speaker's biological endowment as well as of the speaker's cognitive and emotional state. Compared with language, information conveyed by paralinguistic cues is less controllable by the speaker. For example,

TABLE I
Acoustic, Perceptual, and Functional Characteristics of Vocal Behaviors

	Acoustic level	*Perceptual level*	*Functional level*
Voice	Frequency- and amplitude-based parameters	Tone of voice	Voice characteristics
	Mean fundamental frequency (F_0)	Pitch level	High voice
	F_0 variability, F_0 range	Pitch variation, pitch range	
	F_0 contour	Intonation, speech melody	Exaggerated intonation
	Energy level	Loudness	Soft voice, powerful voice
	Measures of frequency spectrum	Timbre, voice quality	Harsh voice, pleasant voice
Speech	Time-based parameters	Speech style	Speech style
	Duration of speech utterances	Amount of speaking	
	Number and duration of silences	Number and length of pauses	
	Sound–silence ratio	Speech rate, tempo	Hectic speaking, halting speaking
		Filled pauses	
		Intrusions	
		Disfluencies	Hesitant speaking
Conversation	Time-based parameters	Conversation style	Conversation style
	Onset and offset of speech passages	Opening and closing of conversation	
	Speaker switching	Speaker switching	
		Turn-taking signals	
		Interruptions	Cooperative, intrusive
		Feedback signals	Attentive, inattentive, criticizing

Note. F_0, mean fundamental frequency.

the experience of stress often constricts the vocal cords, thereby creating an increase in vocal pitch. This unintentionally given information has been called "nonverbal leakage." On the other hand, many aspects of voice, speech style, and face and body movement are managed and applied deliberately. For example, a speaker may increase loudness and raise pitch to prevent an interruption by the listener. In that sense, paralinguistic behavior may serve communicative purposes similar to those of verbal behavior. Yet there are some important differences. First, the information is less explicit because there is no unambiguous relationship between the cues emitted and their meanings. Second, compared with language, the relationship between cues and meanings is more iconic than arbitrary (i.e., the cues bear some intrinsic resemblance to the meanings that are expressed). For example, a loud voice by itself expresses a kind of potency. Third, compared with language, paralinguistic signals can be encoded continuously instead of discretely; that is, modifications can be applied in a gradually varying manner without being subjected to fixed steps. For example, the voice can be raised in varying degrees, whereas a verbal appeal requires decisions about the use of specific terms. Implicitness and ambiguity sometimes may be an advantage for the speaker as well as for social interaction. For example, it might be more polite and less disruptive for social harmony to express anger toward the addressee through one's voice or facial expressions rather than through words.

3. THE RELEVANCE OF CULTURE

Culture is not opposed to nature; rather, as a consequence of phylogenetic evolution, it is part of the natural endowment of humans. However, what can vary is

that there are specific forms of culture. In this sense, the term "culture" is used as a set of guidelines that individuals share as members of a particular society and that tell them how to view the world and how to behave in relation to other people. Cultural guidelines are transmitted from one generation to the next mainly by symbols. The most important system of symbols by which culture is transmitted is language. Although the general ability to use language must be viewed as species-specific, the specific language spoken by an individual is not. It is, as Hofstede stated, a "vehicle of culture." In some sense, it is even a part of culture given that it has evolved simultaneously to the cultural tradition. Similar to the norms and rules of culture that are transmitted by language, language itself must be acquired by an individual during ontogenetic development. Therefore, because paralinguistic behavior occurs simultaneously with speech behavior, the question arises as to what extent it can be considered independent of the specific language that an individual uses. Whether such an abstraction is possible must be demonstrated by analyses showing equivalences in paralinguistic behaviors among users of various languages.

If cultural differences are to be investigated as antecedents of individual behavior, meaningful dimensions that differentiate cultures must be specified. The most important approach for doing so was used by Hofstede. Based on factor-analytic evaluation of questionnaires administered to members of more than 50 countries, he extracted four bipolar dimensions that he called "individualism–collectivism," "power distance," "masculinity–femininity," and "avoidance of uncertainty." The most investigated of these four dimensions is individualism–collectivism. Individualist cultures foster the needs, wishes, and desires of individuals; as such, they encourage autonomy, separateness, and uniqueness. Collectivist cultures foster the needs, wishes, and desires of in-groups over individuals; these cultures encourage values such as social harmony, cohesion, cooperation, and conformity. In individualist cultures, the self–other relationship is perceived as more independent, whereas in collectivist cultures, it is construed as more interdependent. Sometimes, collectivist cultures are also characterized as "high-context" cultures, as opposed to the "low-context" cultures typical of the individualist orientation. A high-context culture is one in which much of the information is implicitly inferred from the situational context and from shared experiences. A low-context culture is one in which most of the information is explicitly stated in the verbal message.

4. EVOLUTIONARY CONTINUITY IN COMMUNICATION

If human language communication is compared with primate vocal communication, some qualitative differences can be found. First, the relationship between sound and meaning is arbitrary and not iconic. Second, the minimal sound units (phonemes) have no meaning by themselves but distinguish different meanings (e.g., the phonemes [r] and [l] have no meaning in English but discriminate the meanings of the words "read" and "lead"). Third, the most dominant principle of sound production is articulation and not vocalization. Fourth, single units are not just repeated or intensified but are combined by rules at different hierarchical levels (i.e., phonological, morphological, lexical, syntactic, and semantic).

Although the symbol system of language must be considered a species-specific feature of humans, vocal communication has its precursors in nonhuman animal behavior. The qualitative differences between human language and nonhuman animal vocal communication do not rule out that there also exist evolutionary continuities of communication from animals to humans in forms and functions. As noted previously, paralinguistic behaviors show at least two properties similar to those of animal communication: iconic relationship to meaning and continuous encoding. Therefore, it makes good sense to study the evolutionary roots of language. The communication of sender identity, emotional states, and interpersonal relationships seems to be particularly well suited for study. For example, birds and mammals use harsh, relatively low-frequency sounds when they are hostile and use higher frequency, more tone-like sounds when they are frightened, appeasing, or approaching in a friendly manner. In 2002, Soltis and colleagues reported that certain responses of squirrel monkeys to audiotaped vocal calls reflect the affiliate relationship between sender and receiver. Although Darwin assumed that it serves the adaptation of an organism to communicate its state as precisely as possible to other organisms of the same species to allow cooperation and minimize the danger of lethal conflict, this assumption has been challenged by modern sociobiologists. According to their view, the basis for evolutionary adaptation is not cooperation but rather struggle between individuals, and natural selection favors individuals who successfully manipulate the behavior of other individuals. Applied to paralinguistic behavior, this means that expressive displays serve strategic purposes rather than simply mirroring emotional states and goals.

Based on those assumptions and results, it can be fruitful to look at similarities and continuities in animal and human communication. It is not impossible to find instances of communication that do not obey the rules of the system of language but reflect phylogenetic older layers of communication in humans. For example, some researchers have claimed that the vocalizations of infants are strikingly similar for humans, lions, polar bears, cats, and gorillas.

5. PARADIGMS FOR THE STUDY OF PARALINGUISTIC BEHAVIORS

5.1. A Functional Framework

Paralinguistic behaviors have been investigated in various disciplines. Among them are linguistics, psychotherapy, social psychology, evolution biology, and expression psychology. To evaluate results and hypotheses with respect to cross-cultural issues, a systematic framework is required. Research may be organized around a classification of material aspects (e.g., acoustic or perceptual properties) as well as around functional issues (e.g., questions closely related to social interaction processes). It seems more parsimonious to focus on the functional perspective because in that case the significance of results, in terms of human universals and cultural differences, can be evaluated easier. A basic assumption of the functional approach is that different patterns of behavior (in terms of material properties) may serve the same function in different cultures, and in turn the same pattern of behavior may serve different functions in different cultures. Starting from a functional classification, researchers must determine to what degree a given function is present in different cultures and by what material properties it is realized.

Two different functional paradigms are presented here: the communication model proposed by Schulz von Thun in 1999 and the Brunswikian lens model proposed by Scherer in 1982. The former is orientated toward verbal communication, whereas the latter was originally developed in the context of general perception and later adapted to social perception research.

5.2. The Communication Model of Schulz von Thun

Schulz von Thun's communication model represents an elaboration of Bühler's organon model, the speech act theory, and Watzlawick and colleagues' communication theory. According to that model, the communication process consists of three distinct components: the sender, the message, and the receiver. The sender and receiver do not act completely independent of one another; they may change their roles. The receiver may give feedback signals that affect the sender. The message, which includes verbal as well as nonverbal properties, is emitted by the sender and decoded by the receiver. There is no one-to-one correspondence between what the sender emits and what the receiver decodes; the sender may emit more than the receiver can decode, and the receiver may perceive or infer other characteristics than those emitted by the sender. Concerning the message or, more specifically, the speech utterance, a distinction is drawn between what is said and how it is said. The former refers to the content of the message, whereas the latter refers to the relationship between the sender and the receiver. The relationship is subdivided into three functional aspects: expression of situational and habitual states of the sender, appeal to the receiver, and the kind of interpersonal relationship between the sender and the receiver.

The four functional aspects may be encoded by all material properties, although to different degrees. For example, an appeal to the receiver can be expressed through words underlining the urgency or legitimacy of a request or through characteristics of the voice signaling that the sender does not expect any resistance. Not all aspects are equally present in all messages; the relative weight of each aspect may vary in different situations. One can assume that cultural norms regulate which aspect is displayed to what degree as well as what verbal and nonverbal properties are to be used in which situation.

5.3. The Brunswikian Lens Model

Schulz von Thun's communication model focuses on products or outcomes of communication and does not give much attention to the processes leading to those products. To throw light on these processes, some investigators have adopted Brunswik's lens model as their research paradigm. Two kinds of processes are distinguished: (a) the externalization of the speaker's traits and states leading to objectively measurable or subjectively perceivable cues of a trait or state and (b) the listener's use of different cues to make judgments on the speaker's states or traits. Implementing the lens model involves three steps. One step is to investigate the process of externalization by measuring various physical qualities, perceptual qualities, or both and

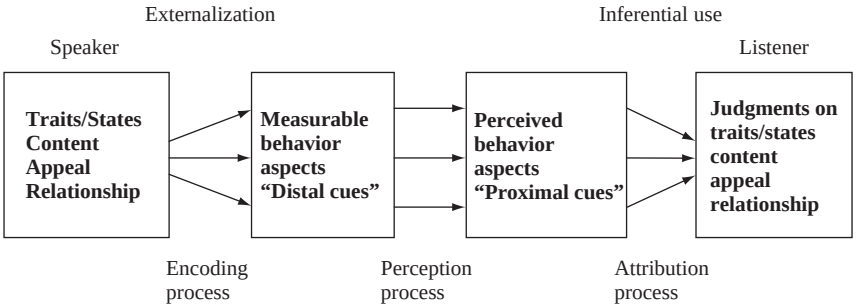


FIGURE 1 A modified version of the Brunswikian lens model.

determining which of these correlate with various speaker features that have been either manipulated or assessed by some objective index (e.g., test scores). This step serves to identify the ecological validity of various distal (i.e., objectively measurable) or proximal (i.e., perceivable) cues. A second step is to determine the accuracy with which listeners can identify the speaker's traits and states from paralinguistic cues. This is done by correlating listeners' judgments with the objective indexes of the speaker's traits and states. A third step is to investigate the specific kind of paralinguistic cues that listeners use to make their attributions. This is performed by correlating various physical or perceptual cues with listeners' judgments. This step identifies listeners' inferential use of cues.

A shortcoming of the lens model is that the communicative function of speech utterances is nearly completely ignored. In terms of Schulz von Thun's communication model, the expression function is stressed without taking into account the aspects of appeal, the kind of interpersonal relationship, and the content of the message. To overcome this weakness, a modified version of the lens model is presented (Fig. 1).

It is assumed that both externalization and use of cues are affected by cultural norms and rules. The model provides a guideline to examine these cultural variations.

6. EXTERNALIZATION OF PARALINGUISTIC BEHAVIORS

6.1. Gender and Age of the Speaker

There is little doubt that there are voice and speech characteristics that mark the personal identity, gender, age, and ethnic background of a speaker. Such long-term speaker-characterizing features are of two different sorts. One type arises from physical characteristics of the vocal and respiratory apparatus and is, to a high degree, beyond volitional control. For example, men and boys usually have louder and lower pitched voices than do women and girls as a result of having a larger respiratory volume as well as longer and thicker vocal cords. Elderly individuals often experience changes in voice quality and loudness due to factors such as ossification of the vocal cords, diminished respiratory capacity, and reduced coordination within the central nervous system. The second type of characterizing features is the product of the way in which the speaker habitually sets his or her vocal apparatus for speaking. These features are under potential control, yet they are not usually deliberately managed by the speaker.

In all societies, age and gender are important social categories, and individuals of different genders and ages may be discernible on the basis of specific voice and speech characteristics. However, both the specific features and the magnitude of differences may be partly affected by cultural factors.

Usually, women and girls show higher pitch and greater pitch variability than do men and boys; however, the magnitude of female–male differences varies across cultures and even across social classes. It may be hypothesized that the magnitude of these differences relates to differences in the masculinity–femininity dimension. Because masculine countries stress gender role differences, voice differences should be more pronounced than in more feminine cultures. Consistent with this hypothesis is that in Japan, a country with a high masculinity index, the difference in pitch (as well as in verbal style) has been found to be higher than that in England and The Netherlands, two countries with a moderate masculinity index. However, the three

countries differ even more with respect to power distance, so that one cannot unambiguously attribute the voice and speech differences to differences in masculinity. It would be of interest to compare various countries of equal power distance differing exclusively in masculinity (e.g., with Japan as a masculine country and Thailand as a feminine country).

In a number of studies, F_0 has been investigated as a function of age. The average F_0 in babies has been found to be extremely high when they are crying (approximately 400–600 Hz) and somewhat lower when they are babbling (approximately 300–500 Hz). It decreases for both males and females with increasing age, at least until adulthood is attained. The changes during adulthood show a marked difference between the sexes, at least in Western societies. For women, no evidence of any systematic changes has been found. On the other hand, average F_0 in men appears to decrease slightly from young adulthood to middle adulthood and then to shift upward from approximately 65 years of age onward. One explanation relates the upward shift in old age to a weakening of the gonads in secreting hormones, a process comparable to puberty, a period when one generally finds a marked decrease in F_0. However, it is also possible to explain this shift without recourse to biology. F_0 has been shown to be affected by both habitual and transitory stress; most people tend to respond to stress with an increase in mean F_0. Thus, it seems reasonable to hypothesize that the upward shift reflects not only age-related physiological changes but also emotional tension caused by factors such as forced retirement, fewer social contacts, decreasing self-sufficiency, and low status associated with old age in Western societies. Cross-cultural comparisons might well shed some light on this issue, but there has been a lack of such studies so far. One may assume, for example, that factors such as the evaluation of old age, in terms of status and institutional support in a particular culture, do play an important role in mediating voice changes.

6.2. Personality Traits of the Speaker

Many attempts have been made to show that paralinguistic behaviors express long-term personality traits. In most of the studies, the central issue is the accuracy with which listeners can judge the speaker's personality on the basis of his or her voice. According to the lens model, however, the first step must be to investigate whether or not personality traits are externalized in objectively measurable distal cues. This can be done by assessing the actual correlations between objectively measured characteristics and valid criteria of personality traits.

There is some evidence from research done in Western countries that personality traits such as extraversion and dominance are indicated by acoustical cues such as energy level and F_0. However, the direction and magnitude of correlations, especially with respect to F_0, turned out to depend on particular situations. In situations where competitive achievement is encouraged, dominant and extraverted people tend to speak in a higher pitched voice (i.e., with a higher mean F_0) than do submissive and introverted people. However, in neutral situations or in situations where communicative abilities are stressed, extraversion and dominance are associated with a lower pitch. Presumably, different facets of extraversion are addressed. Achievement-orientated situations favor ergotropic facets of energy, ascendance, and ambition, whereas neutral or interpersonal communication situations evoke a preponderance of trophotropic (rest) facets of warmth and gregariousness. Moreover, there are considerable cultural differences even within different Western countries. For example, the correlations are more pronounced in Americans than in Germans.

One may assume that in collectivist countries, the correlations between voice and personality are even smaller. When group conformity is stressed instead of individuality, personality traits as markers of the self should be externalized to a much lower degree than in individualist cultures, where the independent self is more likely exhibit cues that clearly mark individuality and serve self-presentation. Thus, it might not be surprising that the correlations between personality and voice obtained in Germany are smaller than those obtained in the United States given that Germany showed a lower index of individualism than did the United States.

6.3. Situational States of the Speaker

Often, the voice is considered to be a direct means of expression of perceived stress. Hence, there have been numerous attempts to find paralinguistic correlates of situational stress. The significance of acoustic cues has

been tested by comparing the differences against some baseline conditions such as neutral or positive situations. Stress has been experimentally induced by threatening movies, increased cognitive workload, and negative feedback.

Stress is often accompanied by increased muscle tension. Thus, it has been hypothesized that increased tension in the respiratory muscles leads to louder speech and that increased tension in laryngeal muscles leads to higher F_0. Increased tension in the respiratory muscles increases subglottal pressure, thereby producing a slightly higher F_0. Because of the increased muscle tension, the F_0 range of the voice should be narrowed and the number of harmonics should also decrease. Moreover, it has been hypothesized that the speech rate and the number of disfluencies increase with higher stress. In general, these hypotheses have been confirmed; most people respond to increased stress with an increase in F_0 relative to baseline, a decrease in the F_0 range, an increase in the speech rate, and an increase in the number of disfluencies. However, there also are considerable interindividual differences. Some people even show a reversed pattern, that is, a decrease in the mean F_0 and an increase in the F_0 range. There is some evidence that these interindividual differences mirror different coping styles such as repression–sensitization. Although they refer to personality differences in arousal control, there is no doubt that cultural norms strongly affect the expression of arousal by so-called display rules. With respect to facial expression of fear, it could be documented that the displays differ across cultures depending on the social situation. When alone, Japanese and American participants displayed exactly the same facial expressions of fear; however, when in the presence of another person, their expressions differed dramatically, with the Japanese invariably smiling rather than displaying their fear. Because articulatory and facial muscles are partially identical, one may expect some parallels to facial expression in the voice. Developmental studies suggest that the regulation of socially disruptive emotions is acquired at an early age. For example, Japanese preschool children respond to frustrating situations with less aggression and anger (as rated on the basis of voice, face, and body) than do U.S. children. To account for the differences, it is hard to determine whether they reflect differences in open expression or differences in experience of emotions due to different interpretations of frustrating events.

6.4. Emotional Content

A person would not be accepted as speaking a language properly if he or she did not deliver the utterances in the pitch pattern, stress pattern, and temporal pattern of grouping and pausing appropriate for that language. Although these prosodic features affect the linguistic meaning of a sentence, the emotional content of an utterance may be expressed independently. There is a long history of researchers trying to identify specific paralinguistic patterns associated with specific emotional contents, including the contributions of the German "expression psychology" and research by evolution biologists. It is a basic assumption that paralinguistic features exist and that they have an isomorphic relationship to their meaning. The validity of distal and proximal cues indicating a particular meaning can be tested by asking speakers to deliberately express a specific emotion, typically one of the six or seven basic emotions: anger, disgust, fear, happiness, sadness/grief, surprise, and (sometimes) contempt. If physically measurable or proximally perceivable correlates of the expressed meanings can be established, the validity may be ensured. Consistent correlations have been found only for anger and sadness/grief, two emotions that may be characterized as the poles of the dimension of activity. Anger seems to be indicated by high F_0, high F_0 variability, high energy level, and high speech rate. The opposite characteristics seem to be associated with sadness/grief: low F_0, narrow F_0 range, low energy level, and low speech rate. So far, there is little evidence for universal patterns discriminating between the poles of other emotional dimensions such as valence (positive vs negative). One reason might be that a particular emotion or emotional dimension is indicated by specific combinations of different cues (i.e., amplitude, frequency, and temporal patterns) that are difficult to quantify. Another reason might be that culture-specific display rules determine in which situations certain patterns may undergo a so-called ritualization, indicated by exaggeration and repetition of single-pattern elements.

6.5. Appeal to the Listener

Although speech behavior may occur without reference to listeners, in most natural speech situations the speaker addresses listeners with the purpose of exerting some influence on the listeners' thoughts, feelings, and behaviors. This function of speech behavior is often referred to as "appeal." It implies an adaptation

toward the specific addressees. The listener variables to which the speaker may adapt can be of different kinds. For example, the speaker may talk to persons of higher, lower, or equal status; to persons of the opposite sex; to familiar persons, or to strangers. Within empirical studies conducted in this field, the most salient feature of listeners that speakers adapt to is age. Thus, a speaker may talk in a loud voice and very slowly to older people, assuming that they are hard of hearing and have difficulty in comprehending. The best-known phenomenon in this area is so-called baby talk (Ammensprache, Motherese, or infant-directed speech), a voice and speech style used by adults, and even by children, when speaking to infants. Although the phenomenon has been found in a variety of cultures and so may assume the status of a universal, the extent of use and the upper limit of the addressee's age vary from culture to culture. Evidence from diverse linguistic and cultural backgrounds shows that infants prefer the baby talk to the adult speech style by showing more attention, more affective behavior, and more positive vocalizations. Infant-directed speech differs from adult-directed speech by phonetic, morphological, syntactic, and paralinguistic features. Most salient are the paralinguistic features. The most notable acoustic features of infant-directed speech, as compared with adult-directed speech, are its higher F_0, wider F_0 range, slower tempo, and longer pauses. The differences in F_0 between infant-directed speech and adult-directed speech seem to be generally more exaggerated in Western cultures (e.g., United States, Australia, England, Italy, Germany) than in Eastern cultures (e.g., China, Japan, Thailand). This may be partly due to the differences between tonal and nontonal languages (e.g., Mandarin Chinese and Thai vs English, German, and Italian). For example, in tonal languages, overly exaggerated F_0 contours superimposed on a sentence may disrupt the lexical information conveyed by word tone. However, one cannot rule out the possibility that the differences also map onto cultural differences in the acceptability of open emotional expression. For example, in Japanese (a nontonal language), the exaggeration of F_0 contours in infant-directed speech, compared with that in adult-directed speech, is also less pronounced. There appear to be three main functions of infant-directed speech in that (a) it serves to initiate and maintain the child's attention, (b) it communicates empathy, and (c) it facilitates language acquisition. Because language-specific features, such as prosodic, lexical, and syntactic features, increase with increasing age of the child, it can be concluded that facilitating infant socialization by engaging attention and empathy dominates at younger ages, whereas facilitating language acquisition prevails during later infancy.

A specific kind of appeal is used in the context of street traffic. For example, a pedestrian crossing a road must ensure that an approaching car will permit him or her to cross. In this case, eye contact and hand and body movements exchanged between the driver and the pedestrian have the function of emblems. Although these emblems may have evolved from universally valid iconic gestures, they have been culturally transformed and their situation-specific meanings must be learned. Eye contact between the driver and the pedestrian may universally signal "I see you," but in some cultures the action-relevant message is "Okay, cross the street," whereas in other cultures the appeal that is expressed is "Pay attention or it will be dangerous for you." Thus, merely decoding the literal meaning of the gesture will be misleading and may have severe consequences.

6.6. Interpersonal Relationship between Speaker and Listener

When two or more people are engaged in conversation, they take turns speaking. Usually, they manage to achieve a fairly smooth synchronization of utterances without too many interruptions and silences. The emerging "rhythms of dialogue" or "accommodation" patterns are highly dependent on the particular speech situation (e.g., formal vs informal, face-to-face vs telephone conversation) and especially on the kind of relationship between the speakers (e.g., mother–child, teacher–student, inferior–superior, like–dislike). Although considerable variations do exist within a culture due to interindividual differences (e.g., extraversion), cross-cultural differences appear to be even larger. For example, the physical duration of silence following an utterance by one of the communication partners that is experienced as appropriate, polite, or uncomfortable is highly regulated by cultural norms and conventions. Roughly speaking, utterances tend to be longer and interspeaker silences tend to be shorter in Western cultures than in Eastern cultures. This may reflect the difference between low-context cultures and high-context cultures. To a member of a high-context culture, what is omitted is an essential part of the communication. Therefore, silence is highly valued. To a member of a low-context culture, what is omitted

creates ambiguity that will be eliminated by explicit verbalizing.

The coordination between the speaker and the listener is perfect when the speaker sends the correct signals for a change of turn and the listener responds to these signals. Also, the listener may interrupt the speaker to take a turn. Interruptions imposed by the listener may signal positive feedback or act as a power device to get control of the conversation. In this vein, two main types of interruptions are distinguished: cooperative and intrusive. An example of the former is a vocalization signaling agreement, whereas an example of the latter is a loud voice abruptly interrupting the speaker. It has been hypothesized that in a collectivist culture that stresses social harmony, individuals send more cooperative interlocutor signals, whereas in an individualist culture that encourages self-assertion and competition, intrusive interlocutor signals dominate. In 2001, Li tested this hypothesis using simulated doctor–patient conversations between either Canadian or Chinese participants. The comparison of the two intracultural conversation conditions showed that Chinese participants engaged in significantly more cooperative interruptions than intrusive ones, regardless of their roles as doctors or patients. In contrast, the Canadian doctors performed significantly more intrusive interruptions than cooperative ones. Although the coding procedures did not differentiate between linguistic and paralinguistic cues, it can be assumed that the latter determined categorizing to a considerable extent. Researchers have observed similar differences in interruption patterns between members of other collectivist cultures (Japanese and Thai dyads) and members of another individualist culture (British dyads). These findings may be taken as providing strong support for Markus and Kitayama's theory that collectivists construe the self–other relationship interdependently, whereas individualists clearly mark their personal boundaries.

In addition to sending audible feedback signals, interlocutors coordinate their conversations through spatial behavior, eye contact, and touch. Cultures develop more or less explicit rules about the appropriate use of these behaviors. Those who stand or sit too far away are seen as cold, aloof, and withdrawn, whereas those who come too close are seen as embarrassingly intrusive and overly intimate. Anthropologists distinguish between contact cultures and noncontact cultures. People from contact cultures stand closer, face more directly, touch, and look more than do people from noncontact cultures. Contact cultures include Arabs, Latin Americans, Southern Europeans (e.g., Greeks, Turks), and a number of African cultures, whereas noncontact cultures include Asians and Northern Europeans. Although it has generally been observed that U.S. Americans stand less close and touch less than do Arabs and South Americans, there are considerable differences within the United States: Hispanics stand closer than do Anglos, and most studies find that African Americans interact at greater distances, look less, and adopt a less direct orientation than do Whites. Moreover, these differences vary with age; African American children stand closer than do Whites, whereas adolescents and adults stand farther away from one another. Age differences have also been observed in Japan, which is generally regarded as a noncontact culture: in public situations, whereas adults show very little body contact, young Japanese women touch each other often.

7. ACCURACY OF LISTENERS' INFERENCES

7.1. Methodological Considerations

To study the accuracy with which listeners can infer traits or states of the speaker and emotional contents of the message on the basis of the speaking voice, listeners are presented with a voice sample and must make a judgment. The accuracy is assessed by correlating the judgments with some objective criteria of states, traits, or content. The voice samples used as stimuli may be, to a great extent, similar to those elicited in externalization studies. However, to prevent listeners from being biased by the linguistic content of the message, either the stimuli must consist of a standard text (spoken or read) or the linguistic content must be masked. Most masking techniques applied are low-pass filtering and random splicing. Low-pass filtering preserves pitch and loudness contours as well as temporal information, such as pauses and speech rate, but it removes most features of voice timbre due to the elimination of upper frequency parts of the voice spectrum. Random splicing is carried out by splitting the speech signal into short segments and then concatenating the segments in random order. It preserves the static voice characteristics of pitch, loudness, and timbre but eliminates the dynamic characteristics such as pitch contour and temporal properties. Judgments are obtained by having listeners either rate a particular trait, state,

or emotion on a rating scale or choose an item from a list of response alternatives.

7.2. Personality Judgments

In many of the earlier studies investigating the accuracy of personality judgments on the basis of voice, researchers have found that there is consensual agreement among the judges but that the validity of judgment is low. Based on these findings, it has been concluded that the judgments reflect "vocal stereotypes" rather than accurate inferences. This conclusion was challenged by arguing that invalid personality measures have been used and that the functional perspective of perceived personality traits has been neglected. In this vein, it has been suggested that people should be most attuned to those traits about which correct or incorrect identification has implications for actions relevant for survival and reproduction in the evolutionary past (e.g., social dominance, extraversion). More recent studies have shown that these traits can be inferred from voice with at least moderate accuracy. This does not necessarily contradict the situation specificity of distal cues indicating these traits given that judges may take into account the particular speech situation when making inferences. However, similar accuracy scores might not be obtained universally. Cultural differences may emerge due to several factors. First, real differences may exist in mean levels of personality traits such as extraversion and dominance. Second, the measurement of personality traits through questionnaires might not be equivalent across cultures due to various self-accessibility of personality traits and various response styles. Third, personality traits as markers of personality may be externalized to various extents due to cultural rules.

7.3. Judgment of Emotional Content

Considering that the emotional content of an utterance can be expressed independently of lexical and syntactic structure, a question arises as to whether listeners are able to accurately recognize the particular emotion the speaker has portrayed. The criterion for accuracy typically used is whether the judges correctly identify the emotion that the speaker has been asked to portray. Psychologists have often debated whether the recognition of emotions is universal or whether it varies by culture. Three competing hypotheses have been proposed. The strong universalist position follows Darwin and states that the expression and recognition of emotions is biologically programmed and, hence, universal. The weak universalist hypothesis asserts that specific emotional categories are largely culturally specific but that broad dimensions, such as activity/arousal and valence, are universal. The interactionist position acknowledges that emotions are programmed biologically but claims that the process of learning to control both expression and perception is highly dependent on cultural factors.

Empirical studies have yielded the following results:

1. Emotions are recognized from voice and face at a better-than-chance level within and across cultures. In addition, accuracy differences among various emotions are highly similar within and across cultures.
2. Cross-culturally, some emotions are better recognized than others. Among the seven basic emotions, fear and disgust are the most poorly recognized, whereas happiness is the most accurately recognized.
3. Accuracy is higher when emotions were both portrayed and identified by members of the same national, ethnic, or regional group, suggesting a so-called in-group advantage.
4. Cross-cultural accuracy is different across channels. In general, the face seems to be the best source of information, although there are emotion-specific differences. Happiness is most accurately recognized from the face but least accurately recognized from the voice. In comparison, anger and sadness are most accurately recognized from the voice but relatively less accurately recognized from the face. Likewise, the in-group advantage seems to be smallest for happiness in the face and greatest for happiness in the voice. Fear shows a large in-group advantage in both the face and the voice, whereas anger shows a relatively small in-group advantage in both the face and the voice.
5. When emotions are expressed by speakers of a foreign language, accuracy is higher for judges whose mother tongue is close in linguistic structure to the language used by the speaker than for judges whose mother tongue is completely different.
6. The evidence for differences in culture per se is not clear-cut. Some results suggest that collectivists have difficulties in both expressing and recognizing emotions (e.g., anger, contempt) that might be potentially disruptive to social interaction. However, it is not clear to what extent the results can be attributed to a mismatch between the cultural background of the speaker and the listener.

Evaluating the three hypotheses formulated previously, it seems that the strong universalist position

cannot be held. There is some support for the second hypothesis in that the dimension of activity/arousal, but not the dimension of valence, seems to be present universally. The interactionist hypothesis fits the data best. Certain core components of emotions seem to be universal and are likely to be biological. However, the style of expression and recognition is governed by cultural conventions. Linguistic rules also appear to have a strong impact. Although linguistic content has been controlled experimentally, language-specific prosodic rules are likely to affect both expression and recognition. This parallels the results found in infant-directed speech, where language-specific and language-independent principles also seem to be intricately linked.

There has been some criticism concerning the studies of emotional recognition centered around the fact that both stimuli and elicitation of responses may lack ecological validity. Also, because responses are typically elicited as forced choices, the high accuracy scores obtained merely show that a particular emotion can be reliably discriminated from other emotions. They do not show that a particular emotion can be correctly identified in reality when no response alternatives are provided. In addition, some researchers have questioned the ecological validity of stimuli. Although deliberately portrayed emotions may exhibit stylized forms of paralinguistic behaviors intended to communicate, these might not be typical in the spontaneous expression of experienced emotions. Instead of mirroring actual emotional states, the latter may be influenced by subjective control mechanisms such as coping and interpersonal manipulation strategies. Moreover, the same emotion may produce different types of expression.

Finally, it has been argued that in face-to-face communication, the listener is presented with information from multiple channels. Thus, optical information may override acoustic information. Compared with judgments of the voice alone, the addition of video information—especially of the face—has been found to be associated with a large additional increase in cross-cultural accuracy. Moreover, in natural communication settings, judgments on voice and behavioral responses to voice usually are affected by the linguistic content of the message. In some cases, linguistic and paralinguistic sources of information may lead to similar inferences, but in others there may be discrepancies. The source of information on which the listener relies most may vary according to the particular speech situation, but it may also be affected by cultural factors.

8. USE OF PARALINGUISTIC CUES BY LISTENERS

8.1. Methodological Considerations

To investigate the kind of acoustic cues that listeners use, correlations between inferences and acoustic measures of the speech signal may be applied. However, results might not always be conclusive because one cannot be sure which of the many cues available have been used and to what extent important cues went unmeasured. Thus, it is useful (and eventually necessary) to manipulate the speech signal. Although speakers can be asked to deliberately vary the paralinguistic features, they may unintentionally vary other features as well. Therefore, experimenters must be able to manipulate features synthetically. This can be done either by systematically changing the digitized speech sample with a computer or by using a speech synthesizer. Typically, F_0, energy level, and pauses have been manipulated.

The aim of some researchers is to identify proximal representations that lead to particular inferences. In that case, listeners are asked to judge the speech signal with respect to perceptual attributes such as soft voice, high voice, powerful voice, unpleasant voice, and halting or hectic speaking.

8.2. Cues to Personality Traits and Situational States

Results obtained in studies on acoustic cues leading to inferences of personality traits and situational states are not quite conclusive. Typically, mean energy level and energy variation are positively associated with dominance and extraversion judgments. With respect to F_0 and speech rate, the results are less consistent and seem to be both situation- and gender-specific. Some researchers have found that speech rate is positively correlated with extraversion and dominance, whereas other researchers have found a negative correlation. High mean F_0 seems to lead to the attribution of high dominance in male voices but not in female voices. Situational stress may be inferred from an increased F_0 relative to a baseline.

Results obtained in studies using perceptual cues indicate that extraversion and dominance are inferred from loud and powerful voices. This is consistent with the findings on acoustical cues. Results on inferences drawn from pitch are as inconsistent as those drawn from F_0. High-pitched male voices are sometimes

judged as extraverted and dominant and, at the same time, as emotionally labile and disagreeable, although they are sometimes judged as less dominant and more agreeable. High-pitched female voices are usually judged as kind, humorous, immature, and agreeable.

Two reasons may account for the inconsistent results. The first is that judges rely more on configurations of multiple cues than on single cues. In addition, F_0 and energy level may be acoustically insufficient to describe relevant configurations. The spectral structure of speech, which determines how the voice sounds, may outweigh the importance of F_0 and energy level. For example, depending on the spectral structure, a high-pitched voice may sound childlike and be perceived as friendly, or it may sound shrill and be perceived as angry and aggressive.

The second reason is that judges are guided not only by vocal stereotypes but also by implicit personality theories. These are characterized by overgeneralizations and illusory correlations such as a negative correlation between dominance and emotional stability.

Although systematic cross-cultural studies on cues to attributions of traits and states are rare, there seems to be evidence of cultural differences. Because average loudness and tempo vary from culture to culture, misattributions from persons not familiar with the cultural standard may occur frequently. For example, Anglos who commonly speak in a softer voice and are not familiar with the Latin standard may erroneously attribute dominance and aggression to the louder voices of Latin people. Even familiarity with the general cultural standards does not prevent misattributions given that sometimes it is essential to analyze the situational demands in a culture-specific way. For example, in cultures stressing power distance, a higher ranked person will be judged as competent when speaking in a loud voice and a low tempo, whereas the same cues may lead to the attribution of boldness when exhibited by a subordinate person.

Cultural differences also seem to exist with respect to implicit personality theories. For example, the Chinese *shi gu,* which refers to "someone who is worldly, devoted to his or her family, socially skilful, and somewhat reserved," is not an easily found stereotype in Western countries. On the other hand, the Western concept of an "artistic type" seems absent in Chinese. Thus, although it is easy for speakers of English or Chinese to refer to single personality traits (e.g., socially skilled, creative), each culture generates its own personality syndromes.

8.3. Cues to Emotional Content

Attempts to identify a set of vocal cues that reliably differentiate among the basic emotions have not yet been successful. In a classic experiment in 1972, Uldall presented listeners with a number of sentences on which 16 different F_0 contours had been imposed synthetically. The contours differed with respect to F_0 range, F_0 reached at end of contour, and shape. The listeners were asked to rate each sentence on several bipolar adjective scales representing the three dimensions of emotional meaning—valence, activity, and potency—defined by Osgood and colleagues in 1957. Both high activity and high potency were inferred from contours characterized by a wide F_0 range and a change of direction. A change of direction combined with a final rise led to the impression of a positive valence. The more smooth or monotonous a contour was, the more the judged meaning was dependent on the linguistic content of the utterance. The more lively a contour was, the more the judged meaning was stable and independent of linguistic content.

8.4. Selective Attention to Various Cues

In collectivist cultures, given that direct confrontation of another person is considered rude and undesirable and that paralinguistic information is less direct and less explicit than verbal information, one may assume that collectivist persons are more sensitive to paralinguistic cues than are individualists, who should be more attuned to linguistic information. In 2002, Kitayama and Ishii tested this hypothesis by using an interference paradigm similar to a Stroop task (i.e., a recognition task in which the print color interferes with the lexical meaning of a written word) to assess spontaneous attention to either linguistic or paralinguistic cues. In this modified Stroop task, lexical emotional content of a word interfered with vocal emotional content. Lexical and vocal meanings had previously been checked as equivalent in extremity in the two languages. Japanese participants (members of a collectivist culture) and English participants (members of an individualist culture) had to make a judgment of either vocal emotion or lexical meaning of an emotionally spoken evaluative word. As hypothesized, an interference effect by competing word evaluation in the vocal emotion judgment was significantly stronger in English than in Japanese. In contrast, an interference effect by competing vocal emotion in the word

evaluation judgment was stronger in Japanese than in English. This finding suggests that individualists are more attuned to direct communication of feelings involving words, whereas collectivists are more attuned to less explicit communication involving paralinguistic features. This does not contradict the notion that open expression of feelings is generally less acceptable in collectivist cultures than in individualist cultures; instead, it suggests that the communication of feelings in collectivist cultures involves paralinguistic behaviors rather than linguistic behaviors.

9. CONCLUSION

9.1. Universals and Cultural Differences

The aim of cross-cultural research on paralinguistic behaviors is to investigate to what extent such behaviors are used in a ubiquitously uniform manner and to what extent they are shaped by cultural factors. Addressing this issue, a modified version of the Brunswikian lens model may serve as a functional framework. Two kinds of processes are distinguished: externalization of paralinguistic behaviors by the speaker and use of paralinguistic cues by the listener. Both kinds of processes seem to be characterized by both universal and culture-specific principles.

The most uniform use of paralinguistic behaviors across cultures can be observed in adult speech directed toward infants. When speaking to an infant, adults accommodate their vocal behavior to the infant's cognitive ability and signal empathy, and the infant responds with attentive and affective signals. The paralinguistic features indicating infant-directed speech seem to be universal, although they are less pronounced in Eastern cultures than in Western cultures. The cultural differences cannot be attributed entirely to language differences (nontonal vs tonal languages). They also appear to map onto cultural differences in the social acceptability of open expression of emotion.

There is additional evidence for universality in decoding emotional meaning from paralinguistic behaviors. When emotions are intentionally portrayed, they are recognized quite accurately from voice and face both within and across cultures. This suggests that certain core components of emotional expression are not conventional and are probably biological. However, this does not imply that emotional experience is uniform across cultures or that specific emotions are universally expressed in the same manner and with similar frequency. The wide variations observed in various cultures seem to reflect different display rules, that is, cultural rules that govern which paralinguistic behaviors are permissible and which are inappropriate in a social setting.

The most striking differences among cultures have been observed in paralinguistic behaviors that regulate the interpersonal relationship between speakers and listeners. It seems that members of individualist cultures send more signals that indicate self-focusing tendencies, whereas members of collectivist cultures exhibit more cues that stress harmony and cooperation. When giving feedback, listeners from individualist cultures tend to show their own feelings toward the utterance of the speaker, whereas in collectivistic cultures it is more important for listeners to signal a positive attitude even when they do not feel it.

Cultural differences have also been observed in the extent to which listeners use paralinguistic cues. Being presented with conflicting information from various channels, it seems that collectivists tend to be more sensitive to the less explicit information conveyed by the voice, whereas individualists rely more on the explicit information transmitted by the spoken words. This difference may reflect the difference between high-context cultures, in which the social meaning of situations is processed in a holistic mode, and low-context cultures, in which an analytic processing mode stressing the abstract meaning of words is preferred.

9.2. Applications to Intercultural Encounters

Cultural differences in paralinguistic behaviors have consequences with respect to intercultural encounters. The appropriate use of paralinguistic behaviors is needed to communicate effectively in a wide variety of intercultural situations. Success in business, teaching, politics, and private interpersonal relationships will depend in great part on the understanding and ability to display the culturally appropriate paralinguistic behavior when communicating with persons who are culturally different. Misunderstandings due to violations of cultural rules can result even if the words and grammar of the foreign language are appropriately mastered. In some cultures, the expectation is to mask, attenuate, or neutralize vocal and facial expressions of one's feelings. The listener must

recognize finer nuances of paralinguistic cues and pay more attention to the situational context in these cultures, as compared with cultures where feelings are openly expressed, to draw inferences about the feelings of persons from these cultures. Most notably, self-disclosure on the part of the speaker is likely to be interpreted as crude or even exhibitionistic in these cultures and may cause embarrassment.

On the other hand, in some cultures feelings are supposed to be dramatized, and so their withdrawal or suppression may lead to the impression of a cool and reserved person. However, as a practical maxim, one can assume that the lost benefit resulting from acting too reservedly in a culture stressing self-disclosure will be less than that resulting from acting too overtly in a culture where self-disclosure is taboo. The reason is that people generally pay more attention to behavior that is shown than to behavior that is not shown.

Lack of effectiveness in intercultural encounters may also result from inappropriate listening. The key to effective listening is participation. However, participation is inferred from different paralinguistic feedback cues in different cultures. In individualist cultures, feedback is expected to reflect the listener's comment on and "honest" attitude toward the content of the message conveyed by the speaker (e.g., feedback to the proposal of a business partner). In collectivist cultures, the listener will do better to just signal that he or she is actively attending to the speaker without giving his or her own comments.

Acknowledgments

The author thanks John Berry, Richard Block, and Harry Triandis for valuable comments.

See Also the Following Articles

Emotion ■ Emotion and Culture ■ Interpersonal Behavior and Culture

Further Reading

Argyle, M. (1996). *Bodily communication* (2nd ed.). London: Routledge.

Aronson, E., Wilson, T. D., & Akert, R. M. (1994). *Social psychology: The heart and the mind*. New York: HarperCollins.

Bühler, K. (1987). Das Organon-Modell der Sprache. In K. R. Scherer, A. Stahnke, & P. Winkler (Eds.), *Psychobiologie* (pp. 321–327). Munich, Germany: Deutscher Taschenbuch Verlag.

Cohn, J. F., Zlochower, A. J., Lien, J., & Kanade, T. (1999). Automated face analysis by feature point tracking has high concurrent validity with manual FACS coding. *Psychophysiology*, *36*, 35–43.

Dabbs, J. M., Jr., & Mallinger, A. (1999). High testosterone levels predict low voice pitch among men. *Personality and Individual Differences*, *27*, 801–804.

Darwin, C. (1859). *On the origin of species by means of natural selection*. London: John Murray.

Darwin, C. (1872). *The expression of emotions in man and animals*. London: John Murray.

Dawkins, R. (1994). *The selfish gene* (2nd ed.). Oxford, UK: Oxford University Press.

DeVito, J. A. (1997). *Human communication* (7th ed.). New York: Longman.

Ekman, P., & Friesen, W. V. (1978). *Manual for the facial action code*. Palo Alto, CA: Consulting Psychologists Press.

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin*, *128*, 203–235.

Grammer, K., & Eibl-Eibesfeldt, I. (1993). Emotions-psychologische Aspekte im Kulturenvergleich. In A. Thomas (Ed.), *Kulturvergleichende Psychologie* (pp. 298–322). Göttingen, Germany: Hogrefe.

Helfrich, H., & Wallbott, H. G. (1986). Contributions of the German "expression psychology" to nonverbal behavior research: IV. The voice. *Journal of Nonverbal Behavior*, *10*, 87–204.

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across cultures* (2nd ed.). Thousand Oaks, CA: Sage.

Itakura, H. (2001). *Conventional dominance and gender: A study of Japanese speakers in first and second language context*. Amsterdam: Benjamins.

Jaffe, J., Beebe, B., Feldstein, S., Crown, C., & Jasnow, M. D. (2001). Rhythms of dialogue in infancy: Coordinated timing in development. *Monographs of the Society for Research in Child Development*, *66*, 1–131.

Kitamura, C., Thanavishuth, C., Chulalongkorn U., Burnham, D., & Luksaneeyanawin, S. (2002). Universality and specificity in infant-directed speech: Pitch modifications as a function of infant age and sex in a tonal and non-tonal language. *Infant Behavior and Development*, *24*, 372–392.

Kitayama, S., & Ishii, K. (2002). Word and voice: Spontaneous attention to emotional utterances in two languages. *Cognition and Emotion*, *16*, 29–59.

Li, H. Z. (2001). Cooperative and intrusive interruptions in inter- and intracultural dyadic discourse. *Journal of Language and Social Psychology*, *20*, 259–284.

Lucas, R. E., Diener, E., Grob, A., Suh, E. M., & Shao, L. (2000). Cross-cultural evidence for the fundamental features of extraversion. *Journal of Personality and Social Psychology*, *79*, 452–468.

Markus, H. R., & Kitayama, S. (1998). The cultural psychology of personality. *Journal of Cross-Cultural Psychology*, *29*, 63–87.

Matsumoto, D. (1993). Ethnic differences in affect intensity, emotion judgments, display rule attitudes, and self-reported emotional expression in an American sample. *Motivation and Emotion*, *17*, 107–123.

Matsumoto, D., Consolacion, T., Yamada, H., Suzuki, R., Franklin, B., Paul, S., Ray, R., & Uchida, H. (2002). American–Japanese cultural differences in judgments of emotional expressions of different intensities. *Cognition and Emotion*, *16*, 721–747.

McCrae, R. R. (2000). Trait psychology and the revival of personality and culture studies. *American Behavioral Scientist*, *44*, 10–32.

Mendoza, E., & Carballo, G. (1998). Acoustic analysis of induced vocal stress by means of cognitive workload tasks. *Journal of Voice*, *12*, 263–270.

Mesquita, B., & Frijda, N. H. (1992). Cultural variations in emotions: A review. *Psychological Bulletin*, *112*, 197–204.

Murata, K. (1994). Intrusive or cooperative? A cross-cultural study of interruption. *Journal of Pragmatics*, *21*, 385–400.

Niwano, K., & Sugai, K. (2002). Intonation contour of Japanese maternal infant-directed speech and infant vocal response. *Japanese Journal of Special Education*, *39*, 59–68.

Noor Al-Deen, H. S. (1997). *Cross-cultural communication and aging in the United States*. Mahwah, NJ: Lawrence Erlbaum.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Urbana: University of Illinois Press.

Protopapas, A., & Lieberman, P. (1997). Fundamental frequency of phonation and perceived emotional stress. *Journal of the Acoustical Society of America*, *101*, 2267–2277.

Ramsey, S. (1997). Double vision: Nonverbal behavior east and west. In A. Wolfgang (Ed.), *Nonverbal behavior: Perspectives, applications, intercultural insights* (pp. 139–167). Seattle, WA: Hogrefe & Huber.

Scherer, K. R. (1979). Personality markers in speech. In K. R. Scherer, & H. Giles (Eds.), *Social markers in speech* (pp. 147–209). Cambridge, UK: Cambridge University Press.

Scherer, K. R. (1982). Stimme und Persönlichkeit: Ausdruck und Eindruck. In K. R. Scherer (Ed.), *Vokale Kommunikation* (pp. 188–227). Weinheim, Germany: Beltz.

Scherer, K. (1997). The state of art in vocal communication: A partial view. In A. Wolfgang (Ed.), *Nonverbal behavior: Perspectives, applications, intercultural insights* (pp. 335–349). Seattle, WA: Hogrefe & Huber.

Scherer, K. R., Banse, R., & Wallbott, H. G. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology*, *32*, 76–92.

Schulz von Thun, F. (1999). *Miteinander reden: Psychologie der Kommunikation*. Reinbek, Germany: Rowohlt–Taschenbuch Verlag.

Searle, J. R. (1969). *Speech acts*. Cambridge, UK: Cambridge University Press.

Soltis, J., Bernhards, D., Donkin, H., & Newman, J. D. (2002). Squirrel monkey chuck call: Vocal response to playback chucks based on acoustic structure and affiliative relationship with the caller. *American Journal of Primatology*, *57*, 119–130.

Trainor, L. J., Austin, C. M., & Desjardins, R. N. (2000). Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychological Science*, *11*, 188–194.

Triandis, H. C. (1995). *Individualism and collectivism*. Boulder, CO: Westview.

Triandis, H. C. (2001). Individualism–collectivism and personality. *Journal of Personality*, *69*, 907–924.

Trull, T. J., & Geary, D. C. (1997). Comparison of the Big-Five factor structure across samples of Chinese and American adults. *Journal of Personality Assessment*, *69*, 324–341.

Tusing, K. J., & Dillard, J. P. (2000). The sounds of dominance: Vocal precursors of perceived dominance during interpersonal influence. *Human Communication Research*, *26*, 148–171.

Uldall, E. (1972). Dimensions of meaning in intonation. In D. Bolinger (Ed.), *Intonation* (pp. 250–259). Harmondsworth, UK: Penguin Books.

Van Bezooijen, R. (1995). Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language and Speech*, *38*, 253–265.

Watzlawick, P., Beavin, J. H., & Jackson, D. D. (1968). *Pragmatics of human communication: A study of interactional patterns, pathologies, and paradoxes*. London: Faber & Faber.

Zahn-Waxler, C., Friedman, R. J., Cole, P. M., Ichiro, M., & Hiruma, N. (1996). Japanese and United States preschool children's responses to conflict and distress. *Child Development*, *67*, 2462–2477.

Zebrowitz, L. A., & Collins, M. A. (1997). Accurate social perception at zero acquaintance: The affordances of a Gibsonian approach. *Personality and Social Psychology Review*, *1*, 204–223.

Part-Time Work

Neil Conway
Birkbeck University of London, London, United Kingdom

1. Introduction
2. Defining Part-Time Work
3. Comparing the Work Attitudes and Behaviors of Part-Time and Full-Time Employees
4. Differences among Part-Time Employees
5. Theoretical Approaches for Understanding Differences across Part-Time and Full-Time Work
6. Conclusion
Further Reading

GLOSSARY

contingent work Work done by part-time employees employed only temporarily (e.g., until a specific organizational goal is met or an operational subtask is finished).
core jobs Employment opportunities commonly viewed as permanent full-time contract positions.
peripheral work Any job in which the individual does not have an explicit or implicit contract for long-term employment or steady hours of employment.

Part-time work has generally been defined as whether an employee reports himself or herself as working part-time. Although the part-time workforce is acknowledged as being a heterogeneous group, most research has concentrated on comparing part-time employees with full-time employees across a range of criteria, most notably work attitudes and behaviors. Differences in work attitudes and behaviors are frequently found between part-time and full-time employees, although the findings are inconclusive and very few theoretical frameworks have been put forward to understand or explain empirical differences. Likely explanations for different attitudes and behaviors across the two groups relate to theories of social exchange and motives for accepting employment.

1. INTRODUCTION

During the past 30 years or so, organizations have increasingly turned to part-time employees because they provide greater scheduling flexibility, meet market demands more efficiently, and reduce wage and benefit costs. For example, between 1971 and 1996 in Great Britain, the proportion of part-time employees nearly doubled from 15 to 29% of the workforce, increasing by 2.6 million. The proportion of part-time employees is similar in the United States, where part-time employees constituted 24% of the workforce and numbered more than 22 million employees in 2001, with the proportion doubling during the previous 40 years. Part-time employees also constitute the major proportion of the workforce for entire industries (e.g., service, retail), and part-time work is much more likely to be performed by certain demographic groups (e.g., women, older/younger workers).

Although this large part-time workforce is clearly important in terms of the increasing dependence on its productivity as well as its social implications, the part-time workforce is also interesting from a psychological

©2004 Elsevier Inc.
All rights reserved.

perspective. For example, some researchers have suggested that part-time workers view work differently than do full-time workers; thus, the way in which the jobs of part-time workers are designed, the human resource practices available to them, and the way in which they are managed all should differ from those of full-time employees.

Organizational researchers have largely neglected researching part-time employment. The majority of studies have simply compared part-time employees with full-time employees across a range of attitudinal outcomes. Too many of these studies have been atheoretical in design and simply documented empirical differences that are of limited lasting scientific value. Moreover, important areas of inquiry (e.g., differences among part-time workers) have been underresearched.

2. DEFINING PART-TIME WORK

There are two necessary components to a definition of part-time work. The first concerns the number of hours worked, and the second concerns comparisons with other forms of employment relationships. Fundamentally, part-time work has been defined by the number of hours worked per week. Government workforce data typically define part-time work as less than 30 to 35 hours per week. Part-time work has also been defined by whether the employee reports himself or herself as working part-time. In organizational research, the self-report definition of part-time work status (i.e., whether employed on a part-time or full-time basis) is used in the great majority of cases.

The second definitional issue is more complex and involves the distinction between part-time work and other forms of employment relationships. The key factor here is the extent to which part-time workers can be regarded as "contingent" or "peripheral" workers. Peripheral employment is equivalent, apart from minor differences, to the conceptual notion of contingent employment. These conceptualizations are rooted in core–periphery models of the workforce in which core jobs are commonly viewed as permanent full-time contract positions. Peripheral or contingent work has been defined as any job in which the individual does not have an explicit or implicit contract for long-term employment or one in which the minimum number of hours can vary in a nonsystematic way.

In early studies, part-time employees were synonymous with peripheral employees. However, there is increasing support for the view that only certain types of part-time employees can be regarded as contingent employees. This view holds that permanent part-time employees are misclassified as contingent simply because they are not employed full-time and that permanent part-time employees are more likely to have contract conditions similar to those of core employees. For example, many permanent part-time positions permit long-term employment. Only certain types of part-time employees, such as part-time employees who are also employed on a temporary basis, are likely to fall within the definition of contingent workers. Distinguishing part-time work from other forms of employment is important because working part-time might not necessarily entail a subjective sense of being peripheral.

In conclusion, part-time work is typically defined by whether the employee reports himself or herself as working part-time. A further definitional issue is sensitivity to differences among part-time jobs, such as the permanence of the positions, because such characteristics are likely to be of substantive theoretical importance.

3. COMPARING THE WORK ATTITUDES AND BEHAVIORS OF PART-TIME AND FULL-TIME EMPLOYEES

Over the past 30 years or so, there has been a slow stream of studies examining part-time work where the dominant approach has been to compare the work attitudes and behaviors of part-time employees with those of full-time employees. Only the relationship between work status and job satisfaction and commitment have received any serious level of coverage, although some studies have considered other relationships such as work status and organizational climate, job involvement, and work characteristics.

Empirical research comparing part-time employees with full-time employees consistently finds differences on various work attitudes and behaviors. However, the findings are inconclusive. For example, across studies, part-time employees have been found to report higher, lower, and equivalent levels of job satisfaction compared with full-time employees. When facets of job satisfaction (e.g., satisfaction with pay, satisfaction with supervisor) have been considered, inconsistent findings remain. What these more in-depth studies have shown, however, is that ratings of satisfaction of part-time employees and those of full-time employees

might not act uniformly across facets of satisfaction. For example, part-time employees may report lower satisfaction with job content but higher satisfaction with social aspects of the work. These studies have drawn attention to the fact that nonsignificant differences in overall ratings of job satisfaction across work statuses may mask differences regarding facets of satisfaction.

Similarly, inconsistent findings have emerged from comparisons of the commitment levels of part-time employees and those of full-time employees. If one looks at the much smaller number of studies that have considered the relationship between work status and work-related behavior in the form of employee turnover, absenteeism, and job performance, the findings are again inconclusive.

A few studies have investigated work status as a moderator of the relationships among a range of job-related attitudes and behaviors. This is a potentially important issue because it may indicate whether predictors of attitudes for full-time employees, such as the psychological contract, are of equal utility to understanding part-time attitudes at work. From existing studies, there is little evidence of work status as a moderator, implying that part-time employees will respond in the same manner as will full-time employees in terms of human resource practices, rewards, and management strategies.

The main conclusion to be drawn from this brief review is that although attitudinal and behavioral differences are frequently observed between full-time and part-time employees, these empirical studies yield inconsistent results. Three further points are worth noting. First, findings on whether work status acts consistently across outcomes within the same study are mixed. Some studies have found that part-time employees report higher ratings of both job satisfaction and commitment, whereas other studies have found that part-time employees report higher ratings of job satisfaction but lower ratings of commitment. Collectively, these studies suggest that theories used to explain differences in job satisfaction between part-time and full-time employees might not be appropriate for organizational commitment. Second, some construct measures used in studies comparing part-time employees with full-time employees are designed with full-time employees in mind, raising the question of whether any differences found across work statuses on such measures can be considered as genuine differences on the underlying construct or instead should be deemed as artifacts of the method used. Finally, a narrow range of methods has been used, with cross-sectional self-report questionnaire surveys being adopted nearly exclusively, where the limitations of such methods are well documented.

4. DIFFERENCES AMONG PART-TIME EMPLOYEES

Until the past decade or so, the heterogeneity among the part-time work group was rarely considered. For example, there were no clear theoretically based rationales for including demographic variables in research designs. Demographic variables were used to refer to age, sex, and marital status as well as to conditions such as organizational tenure and education. The control of demographic variables varied considerably across studies, with some studies not reporting the impact of any demographic variables and others using such unique sets of control variables that generalization was difficult.

In 1990, Feldman attempted to reconceptualize the research on part-time work. He argued that there were also major theoretical distinctions within the part-time work group and identified several key dimensions on which part-time work arrangements can vary, including the following:

- Whether the employee is employed on a permanent or temporary basis
- Whether the employee is recruited directly by the company or through an agency
- Whether the employment is year-round or seasonal
- Whether the job is the employee's main job or a second job
- Whether the individual accepted the part-time job voluntarily or involuntarily due to the unavailability of a full-time position

Feldman argued that these dimensions relate to attitudinal and behavioral outcomes. For example, part-time workers are hypothesized to be more likely to quit when they work part-time on an involuntary basis. Furthermore, there are three key demographic groups within the part-time workforce: students, married women with children, and retirees. These groups are indirectly related to attitudinal and behavioral outcomes in that certain demographic groups will gravitate to certain types of part-time work. For example, retirees are more likely to gravitate to voluntary part-time working arrangements as their main jobs, and these jobs are hypothesized as being more satisfying.

Feldman's research has drawn attention to some of the key dimensions of part-time work, but very few studies have explored whether these dimensions explain differences found in attitudes and behaviors within the part-time work group or, indeed, across part-time and full-time employees. In addition to highlighting important distinctions within the part-time work group, these categories draw attention to the need to control for confounding variables when comparing part-time employees with full-time employees. Researchers are now careful to locate their samples in terms of these categories.

Essentially, research in this area is concerned with what being "part-time" actually means. Some of the main dimensions of part-time work appear to have been identified, but these dimensions have received very little empirical attention and do not appear to be exhaustive.

5. THEORETICAL APPROACHES FOR UNDERSTANDING DIFFERENCES ACROSS PART-TIME AND FULL-TIME WORK

It should be noted that research comparing part-time employees with full-time employees has been largely atheoretical in design. When theories have been put forward to explain differences in attitudes and behaviors, mostly partial inclusion and/or frames of reference, they are applied post hoc to rationalize findings. The theories have received minimal attention and have tended to be applied in a rather cursory way.

5.1. "Part-Time" as a Proxy for Other Key Variables

The composition of the part-time workforce is known to differ from that of the full-time workforce in terms of various demographic (e.g., gender), organizational (e.g., tenure), and labor market (e.g., temporary/permanent) variables. Researchers have argued that these differences explain different attitudes and behaviors found across part-time and full-time employees and that once these differences are controlled for work status, they will no longer be associated with job attitudes and behaviors. In other words, differences in attitudes and behaviors between part-time and full-time employees are not due to the former's "part-timeness." Rather, they are due to the fact that part-time employees are more likely to come from different demographic, organizational, and/or labor market groups; that is, "part-time" acts as a "proxy" for other key variables that are associated with working part-time.

A major problem with this argument is that proxy explanations constitute neither an adequate theoretical explanation nor an adequate psychological explanation for why there may be differences in attitudes and behaviors between part-time employees and full-time employees, and the use of proxy explanations may mask more meaningful theoretical explanations.

5.2. Partial Inclusion

Part-time employees are sometimes said to be "partially included" in the workplace. By spending less time in the workplace and being more involved in extra-organizational roles, they feel less included in organizations than do full-time employees. A number of limitations to this theory have been documented. The theory has been used in contradictory ways to explain attitudes where, for instance, feeling less included has been used to explain both higher and lower levels of job satisfaction. This theory has never been empirically tested, and it is not clear how the theory could be operationalized given that it is such a poorly elaborated concept.

This theory is perhaps better seen as a broad framework within which better understood theories can be placed. For instance, partial inclusion can be seen as embodying a number of ideas that could be explored separately in their own right, including the psychological contract, theories of the relationship between work and nonwork, and social identity theory.

5.3. Frames of Reference

It has been argued that part-time employees have a "frame of reference" that is different from that of full-timers, where the former evaluate characteristics of their work environment by considering how well they compare with those of another self-selected group. For example, part-time employees may be less satisfied with their pay if it compares unfavorably with that of their full-time counterparts.

This approach has been criticized in a similar way to partial inclusion in that it also has been used in contradictory ways to explain attitudes, is untested, and is too vague to operationalize. Furthermore, frame of reference comparisons can be used in an endless number of ways in terms of to whom, and over what, part-time employees choose to compare themselves.

5.4. Social Exchange and the Psychological Contract

The two exchange-based approaches of the psychological contract and social exchange theory have also been used to explain different attitudes and behaviors across work statuses. It has been argued that there are a number of factors that affect the exchange of contributions for inducements across part-time and full-time employees. For example, part-time employees receive fewer inducements such as benefits, task variety, and opportunities for advancement; part-time employees have lower expectations about what they should get from the organization; and part-time employees are more likely to be subject to "Theory X"-type management. These factors will have the effect of creating a perception of perceived inequity or psychological contract violation across work statuses, and they are more likely to lead to part-time employees developing economic relationships, rather than social exchange relationships, with employers. This, in turn, will affect outcome attitude and behaviors such as job satisfaction and organizational citizenship behavior.

Although some of the assumptions behind the social exchange-type approaches are questionable (e.g., that organizations treat part-time employees less favorably, that part-time employees expect less from their work), these approaches have been empirically tested and found to at least partially explain attitude and behavior differences across part-time and full-time employees.

5.5. Preferred Work Status/Schedule

Employees working their preferred work status (i.e., part-time vs full-time) or schedule (i.e., preferred number of hours per week) are often said to have more positive work attitudes. Previous research has found that a discrepancy between scheduled and preferred work hours for both full-time and part-time employees is related to work attitudes and that preferred work status matters more to part-time employees than to full-time employees.

What is not clear is the precise theoretical explanation for why such a discrepancy affects attitudes and behavior. One possible explanation is that employees involuntarily situated in their work arrangements have lower levels of attachment. An alternative explanation drawn from social exchange theory is that employees working their preferred schedules achieve a better balance between contributions and inducements.

6. CONCLUSION

Despite a long history of comparing the work attitudes and behaviors of part-time employees with those of full-time employees, there have been relatively few studies conducted. Studies frequently find attitudinal and behavioral differences across part-time and full-time employees, but what emerges quite categorically is that findings are not consistent across studies. Many previous studies have been atheoretical in design and, thus, have contributed little to understanding the experience of part-time work. Since the 1990s, research on part-time work has been conducted under a stronger theoretical basis, and two particularly promising approaches to a better understanding of the experience of part-time work are social exchange and motives for accepting part-time employment.

Acknowledgment

The author thanks the Leverhulme Trust for its support in the preparation of this article.

See Also the Following Articles

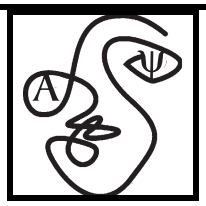
Contracts ■ Cooperation at Work ■ Person–Environment Fit ■ School-to-Work Transition ■ Work Environments ■ Work Teams

Further Reading

Barling, J., & Gallagher, D. (1996). Part-time employment. *International Review of Industrial and Organizational Psychology*, *11*, 243–277.

Conway, N., & Briner, R. B. (2002). Full-time versus part-time employees: Understanding the links between work status, the psychological contract, and attitudes. *Journal of Vocational Behavior*, *61*, 279–301.

Feldman, D. C. (1990). Reconceptualizing the nature and consequences of part-time work. *Academy of Management Review*, *15*(1), 103–112.



Pediatric Psychology

Merritt M. Jensen and Anne E. Kazak
Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA

1. Definitions: Pediatric Psychology and Pediatric Psychologist
2. Pediatric Psychology: Integrating Psychology and Pediatrics
3. Emergence of Pediatric Psychology
4. Growth of Pediatric Psychology within Medical Settings
5. Major Research Areas within Pediatric Psychology
6. Clinical Practice of Pediatric Psychology
7. Pediatric Psychology Training
Further Reading

GLOSSARY

adherence to medical treatment The complex interactions among patients, families, and health care providers in ensuring that treatment regimens are followed as intended.

assessment The evaluation process to describe domains related to physical, intellectual, cognitive, emotional, social, and family functioning.

bereavement The grief process following the loss (typically the death) of a loved one.

childhood chronic illness An ongoing medical condition requiring long-term medical care and follow-up (e.g., sickle cell disease, HIV/AIDS, asthma, cancer, juvenile diabetes); it often affects the life of the affected child and his or her family for an extended time period and can interfere with daily functioning.

cognitive–behavioral therapy (CBT) Psychological intervention that focuses on understanding how maladaptive thought patterns and beliefs lead to behavior that interferes with optimal social, emotional, and behavioral functioning; CBT intervention concentrates on altering thoughts and beliefs so as to change behaviors and emotional states toward a more adaptive perspective.

evidence-based treatment An intervention for which there is empirical support.

interdisciplinary Involving two or more disciplines working toward a common goal; for example, an interdisciplinary team in pediatric oncology might include oncologists, nurses, nutritionists, social workers, physical therapists, occupational therapists, child life specialists, psychologists, and physicians from other medical disciplines.

intervention A systematic approach to facilitate change in a desired direction that may focus on individuals or families (e.g., cognitive–behavioral therapy, family therapy, pain management, consultation).

procedure A medical treatment (e.g., having blood drawn, undergoing a magnetic resonance imaging [MRI] scan).

psychosocial Involves the integration of psychological and social aspects of a person's well-being; sometimes, it is expanded as *biopsychosocial* to include the role of physical (biological) factors.

Pediatric psychology continues to grow and develop as an area of specialty and expertise within clinical psychology. This article highlights the interdisciplinary nature of the field. Research in pediatric psychology, including the emergence of evidence-based interventions, and their use in the clinical practice of pediatric psychology is described. Training requirements for pediatric psychologists are also discussed.

©2004 Elsevier Inc.
All rights reserved.

1. DEFINITIONS: PEDIATRIC PSYCHOLOGY AND PEDIATRIC PSYCHOLOGIST

As described by the Society of Pediatric Psychology, pediatric psychology "is dedicated to research and practice addressing the relationship between children's physical, cognitive, social, and emotional functioning and their physical well-being, including maintenance of health, promotion of positive health behaviors, and treatment of chronic or serious medical conditions." Pediatric psychologists are experts in the areas of developmental, emotional, behavioral, social, and family aspects of acute and chronic illness and injury. They work with children, adolescents, families, and health care teams around the areas of the physical, emotional, social, and cognitive aspects of illness or injury. Within interdisciplinary pediatric health care teams, pediatric psychologists evaluate and intervene to enhance the well-being of children and families.

2. PEDIATRIC PSYCHOLOGY: INTEGRATING PSYCHOLOGY AND PEDIATRICS

The importance of a close working relationship between psychology and pediatrics has long been viewed as essential to providing quality health care for children, adolescents, and their families. In 1965, Jerome Kagan spoke of the need for a "new marriage" between the fields of psychology and pediatrics to meet the needs of children in health care settings. In 1967, Logan Wright published a seminal article delineating the key elements of the field of pediatric psychology, including a call for empirical research supporting clinical practice.

3. EMERGENCE OF PEDIATRIC PSYCHOLOGY

Pediatric psychology is relatively new, having been founded in 1968 as the Society of Pediatric Psychology. Dorthea Ross, Lee Salk, and Logan Wright were early leaders in the field. Interest in working with pediatric patients and their families within medical settings continued to grow during the 1970s. In 1973, the *Journal of Pediatric Psychology* was established and today remains the flagship journal in the field. This journal publishes articles focused on original research, professional practice, and theory in pediatric psychology. Steady increases in membership and growing interest in working with pediatric patients and their families within medical settings occurred during the 1980s and 1990s, culminating in the Society of Pediatric Psychology gaining division status (Division 54) within the American Psychological Association in 1999.

4. GROWTH OF PEDIATRIC PSYCHOLOGY WITHIN MEDICAL SETTINGS

The numbers of pediatric psychologists are increasing within hospitals and other medical settings. Working together with interdisciplinary medical teams, pediatric psychologists serve as consultants and collaborators within inpatient and outpatient hospital settings, primary care and subspecialty clinics, and academic medical settings and schools.

5. MAJOR RESEARCH AREAS WITHIN PEDIATRIC PSYCHOLOGY

Scientific research is an essential part of pediatric psychology. Childhood chronic illness is a major focus of the field, although treatment of acute illness and attention to health promotion and prevention are also emphasized. Designing and evaluating evidence-based treatments and interventions is an important recent focus.

5.1. Chronic Illness

Frequently studied childhood chronic illnesses include (but are not limited to) asthma, cancer, cystic fibrosis, diabetes, HIV/AIDS, juvenile rheumatoid arthritis, organ transplantation, seizure disorders, pain syndromes, neuromuscular diseases, and sickle cell disease. Childhood chronic illness research focuses on how multiple aspects of diseases and treatments affect children, adolescents, and families, including adjustment to and coping with chronic illness, adherence to medical procedures and treatments, pain management, feeding and sleep disturbances, intellectual and cognitive development, social and emotional functioning with peers and family, family adjustment to chronic illness, and promotion of health and quality of life.

5.2. Assessment

Systematic assessment, using standardized psychological tests, can facilitate understanding of the intellectual, cognitive, social, and emotional impacts of illnesses or injuries. The assessment process provides essential information for developing effective treatments and necessary information to health care teams, families, and school personnel. For example, many diseases or illnesses adversely affect the child's central nervous system, causing physical, cognitive, and/or socio-emotional deficits that may influence the child's ability to perform competently at home, with peers, and in school.

5.3. Intervention: Evidence-Based Treatments

Carefully designed and controlled studies evaluating the effectiveness of manual-based interventions with children and their families have increasingly become a part of research and practice in the field of psychology. Beginning in 1999, the *Journal of Pediatric Psychology* published a series of reviews describing evidence-based treatments in pediatric psychology, including interventions for treating recurrent pediatric headache, recurrent abdominal pain, procedure-related pain, disease-related pain, pediatric obesity, severe feeding problems, disease-related symptoms, bedtime refusal and nighttime wakings, nocturnal enuresis, constipation and encopresis, and adherence to treatment regimens.

There are several effective treatments that may be used to treat children's pain. For example, cognitive–behavioral therapy (CBT) is an effective, well-established treatment for procedure-related pain, disease-related pain, and recurrent abdominal pain in children and adolescents. Examples of CBT techniques include teaching the child and family relaxation skills such as the use of deep breathing exercises, using guided imagery, and using distraction such as counting out loud or playing with toys. Active coaching by a caregiver and/or medical team member encourages the child to use self-soothing coping skills, which is another effective CBT intervention for pain. Relaxation and self-hypnosis are useful treatments for recurrent pediatric headaches. Biofeedback, whether alone or combined with relaxation or other CBT treatments such as those described previously, is also promising for headaches. In addition, guided imagery is an effective treatment for children with cancer who are undergoing chemotherapy.

Effective treatments for childhood obesity target several behavioral areas at one time, for example, promoting healthy eating patterns while also increasing physical exercise. For severe feeding problems, rewarding the child when he or she engages in desired behaviors (e.g., eating a particular food) and ignoring undesired responses (e.g., spitting out food) are useful interventions.

Evidence-based interventions for bedtime refusal and other childhood difficulties, including enuresis and encopresis, are promising. Educating parents about healthy sleep hygiene during the prenatal and newborn periods is helpful for bedtime refusal and night wakings in young children. Implementing a positive bedtime routine is another promising intervention. Extinction techniques, including parental ignoring of the child's cries so that the child learns to fall asleep on his or her own, is also effective but is often less feasible given that parents might not be willing to attempt this strategy. For nocturnal enuresis (i.e., repeated nighttime urination in places other than a toilet, e.g., in bed), addressing both the biological and behavioral aspects of the problem is most effective. Using a urine alarm combined with the medication desmopressin is a promising intervention. Fewer effective, evidence-based treatments for constipation and encopresis (i.e., repeated bowel movements in locations other than a toilet, e.g., in underwear) have been identified. The most promising treatments combine medical intervention (e.g., fiber, laxatives) with a behavioral component such as positive reinforcement.

Understanding the most effective ways in which to encourage adherence to medical regimens is an essential area of pediatric psychology research. Increased parental supervision and monitoring of medication use, coupled with increased attention and support from a health care team, improves adherence for children and adolescents with asthma. For individuals with juvenile rheumatoid arthritis, treatments using behavioral strategies (e.g., positive social attention, feedback from the child's family, token reward systems) are useful. Multicomponent treatments focused on teaching problem-solving, negotiation, and decision-making skills are promising for helping children, adolescents, and families to improve adherence to complex regimens for juvenile (type 1) diabetes.

5.4. Prevention

Often, the best therapeutic intervention is to prevent the occurrence of an illness or injury. Pediatric psychologists are actively involved in promoting healthy attitudes

and behaviors in children, adolescents, and families and in working to prevent disease, illness, and injury. Research targeting lifestyle choices and behaviors during childhood and adolescence, including healthy eating, exercise, drug avoidance, injury prevention, and safe sexual practices, encompasses the many prevention topics studied by pediatric psychologists.

5.5. Family Systems

Increasingly, the focus of pediatric psychology has transitioned from focusing exclusively on the child and adolescent to a more systems-focused perspective emphasizing the important roles that parents, caregivers, siblings, relatives, peers, teachers, the health care team, and other community contacts play in the child's functioning in relation to illness or injury. Families facing illness must balance medical management of the disease with feelings of fear, uncertainty, and grief. How successfully families are at navigating this challenge is influenced by the families' pre-illness functioning and adjustment, levels of support and resources, and the medical aspects of the illness. Families who demonstrate strength-based skills, such as effective communication, clear family boundaries, active coping, flexibility, and the ability to balance the demands of the illness with other family responsibilities, are more likely to demonstrate increased well-being throughout course of the illness.

6. CLINICAL PRACTICE OF PEDIATRIC PSYCHOLOGY

6.1. Settings in Which Pediatric Psychologists Work

Many pediatric psychologists are employed in, or have close affiliation with, hospitals. They work in emergency, inpatient, and/or outpatient facilities in freestanding children's hospitals, university-affiliated medical centers, community hospitals, and satellite clinics. A smaller percentage of pediatric psychologists are based in primary care, and others work in private practice serving children and families. Pediatric psychologists are also employed within schools and in other community settings.

6.2. Roles of Pediatric Psychologists

The professional roles of pediatric psychologists are as diverse as the environments in which they work. Pediatric psychologists often serve as consultants to interdisciplinary medical teams within inpatient, outpatient, or primary care settings. Assessment of psychosocial, neuropsychological, and/or neurodevelopmental functioning of children with acute or chronic illness, injury, or developmental delay is another important activity. Pediatric psychologists often work directly with children and their families around the adjustment to acute or chronic illness or injury, including how the illness or injury will affect the children's cognitive, emotional, and social functioning. Adjustment to illness or injury may also include addressing adherence to medical regimens, coping with painful or stressful medical procedures, and exploring how the child's and family's beliefs about the adversity of the illness or injury are affecting their thoughts, feelings, and relationships with others. Bereavement work with children and families facing death due to illness or injury is another important activity.

7. PEDIATRIC PSYCHOLOGY TRAINING

Pediatric psychologists are typically licensed, doctoral-level psychologists. Many pediatric psychologists have backgrounds in clinical psychology, whereas others have training in areas such as developmental or health psychology. All have completed 4 to 5 years of graduate-level coursework and clinical training, a 1-year intensive clinical internship in a hospital or other setting related to health care, and 1 or more years of postdoctoral training.

In addition to didactic and experiential learning in basic psychological principles, pediatric psychologists receive specialized training and instruction in health care issues that affect children, adolescents, and their families. Focused training in pediatric psychology often includes instruction in the areas of adherence and health promotion, management of painful medical procedures, social and emotional aspects of pediatric medicine, and prevention of illness and injury. While training within diverse pediatric health care settings, pediatric psychologists gain advanced experience in working with children and adolescents with a broad range of medical illnesses.

Pediatric psychologists continue to use evidence-based research to inform clinical practice within interdisciplinary care environments. With the growth of this

relatively newly emerging field, the role for pediatric psychologists will clearly broaden and diversify as the needs of changing medical environments dictate.

See Also the Following Articles

Child Development and Culture ■ Diverse Cultures, Dealing with Children and Families from ■ Health and Culture

Further Reading

American Psychological Association. (n.d.). *Division 54: Society of Pediatric Psychology*. www.apa.org/about/division/div54.html.

Drotar, D. (1995). *Consulting with pediatricians: Psychological perspectives*. New York: Plenum.

Drotar, D. (2000). *Promoting adherence to medical treatment in chronic childhood illness: Concepts, methods, and interventions*. Mahwah, NJ: Lawrence Erlbaum.

Holden, E. W., Deichmann, M. M., & Levy, J. D. (1999). Empirically supported treatments in pediatric psychology: Recurrent pediatric headache. *Journal of Pediatric Psychology*, *24*, 91–109.

Janicke, D. M., & Finney, J. W. (1999). Empirically supported treatments in pediatric psychology: Recurrent abdominal pain. *Journal of Pediatric Psychology*, *24*, 115–127.

Jelalian, E., & Saelens, B. E. (1999). Empirically supported treatments in pediatric psychology: Pediatric obesity. *Journal of Pediatric Psychology*, *24*, 223–248.

Journal of Pediatric Psychology (1975–present). Oxford University Press. http://jpepsy.oupjournals.org/.

Kagan, J. (1965). The new marriage: Pediatrics and psychology. *American Journal of Diseases of Childhood*, *110*, 272–278.

Kazak, A. (2002a). *Journal of Pediatric Psychology* (*JPP*), 1998–2002; Editor's vale dictum. *Journal of Pediatric Psychology*, *27*, 653–663.

Kazak, A. (2002b). *Journal of Pediatric Psychology:* A brief history (1969–1999). *Journal of Pediatric Psychology*, *25*, 463–470.

Kazak, A. E., Rourke, M. T., & Crump, T. A. (2003). Families and other systems in pediatric psychology. In M. Roberts (Ed.), *Handbook of pediatric psychology* (3rd ed.). New York: Guilford.

Kerwin, M. E. (1999). Empirically supported treatments in pediatric psychology: Severe feeding problems. *Journal of Pediatric Psychology*, *24*, 193–214.

Lemanek, K. L., Kamps, J., & Brown Chung, N. (2001). Empirically supported treatments in pediatric psychology: Regimen adherence. *Journal of Pediatric Psychology*, *26*, 253–275.

McGrath, M. L, Mellon, M. W., & Murphy, L. (2000). Empirically supported treatments in pediatric psychology: Constipation and encopresis. *Journal of Pediatric Psychology*, *25*, 225–254.

McQuiad, E. L., & Nassau, J. H. (1999). Empirically supported treatments of disease-related symptoms in pediatric psychology: Asthma, diabetes, and cancer. *Journal of Pediatric Psychology*, *24*, 305–328.

Mellon, M. W., & McGrath, M. L. (2000). Empirically supported treatments in pediatric psychology: Nocturnal enuresis. *Journal of Pediatric Psychology*, *25*, 193–214.

Mindell, J. (1999). Empirically supported treatments in pediatric psychology: Bedtime refusal and night wakings in young children. *Journal of Pediatric Psychology*, *24*, 465–481.

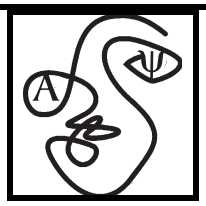
Mullins, L. L., & Chaney, J. M. (2001). Pediatric psychology: Contemporary issues. In C. E. Walker, & M. C. Roberts (Eds.), *Handbook of clinical child psychology* (3rd ed.). New York: John Wiley.

Powers, S. (1999). Empirically supported treatments in pediatric psychology: Procedure-related pain. *Journal of Pediatric Psychology*, *24*, 131–145.

Roberts, M. C. (2003). *Handbook of pediatric psychology* (3rd ed.). New York: Guilford.

Walco, G. A., Sterling, C. M., Conte, P. M., & Engel, R. G. (1999). Empirically supported treatments in pediatric psychology: Disease-related pain. *Journal of Pediatric Psychology*, *24*, 155–167.

Wright, L. (1967). The pediatric psychologist: A role model. *American Psychologist*, *22*, 323–325.



Perception and Culture

J. B. Deręgowski
University of Aberdeen, Aberdeen, United Kingdom

1. Introduction
2. The Notion of Culture
3. Visual Illusions
4. Pictorial Perception
5. Perception of Patterns
6. Eidetic Imagery
7. Studies of Perception of Color
8. Perception of Time
9. Concluding Remarks
Further Reading

GLOSSARY

cross-cultural An imprecise term used to describe comparisons between groups differing in culture but also often differing in other respects (e.g., gene pool, environment).
eidolicity Characteristic of flat figures responsible for perception of illusory depth.
epitomy Characteristic of flat figures responsible for their recognition as representing objects.
horizontal–vertical illusion Misperception of T-shaped and L-shaped figures.
Kwengo calipers A test in which individuals adjust a pair of calipers to an angle presented in perceptually flat or perceptually solid figures.
Muller–Lyer illusion Misperception of the length of lines bounded by arrowheads.

Cross-cultural studies of perception show differences among populations. They contribute more to the understanding of the perceptual processes than do analogous studies of nonhuman species. In spite of this, cross-cultural studies have not flourished greatly and have been concerned mostly with visual perception.

1. INTRODUCTION

Cross-cultural differences have long been the subject of travelers' reports. For example, it was reported that somewhere near the present-day Afghanistan, Alexander the Great's army encountered men who had one foot much larger than the other. These feet, in addition to their usual function, served as parasols when the men lay prone on their backs.

Systematic cross-cultural investigations are relatively recent, although not far temporally removed in time from the beginnings of scientific psychology. The first significant cross-cultural psychological study, the scope of which has never been exceeded, was carried out as part of the Cambridge Anthropological Expedition to the Torres Straits. Several psychologists, the most prominent of whom was W. H. R. Rivers, collected data on several perceptual functions such as olfactory and cutaneous sensations. However, much of the data collected, just like the data most ardently sought by research workers of more recent years, concerned visual perception. Later studies became even narrower in the scope of the phenomena investigated. With the notable exception of Reuning's investigations of the visual abilities of

©2004 Elsevier Inc.
All rights reserved.

the Bushmen, the studies are generally concerned with only one aspect of vision.

2. THE NOTION OF CULTURE

The term "cross-cultural" is vague because there is no definition of culture that would satisfy experimental psychologists, for example, by enabling them to decide whether two individuals come from different cultures with as great an aplomb as the psychologists can decide whether the individuals differ in terms of a perceptual skill or short-term memory. However, it is generally believed that the vectors of genetic endowment and the impact of the natural environment are insufficient to explain behavior and that the manmade components of the environment, such as language, architecture, customary rituals, and social organization, form a third vector, the "cultural" vector.

Therefore, there is a widespread conviction that the vectors that determine individuals' responses are threefold: genetic, environmental, and cultural. However, it is also generally accepted that the vectors are not mutually independent. For example, it is accepted that, when a "hunting and gathering" population changes its mode of life (a cultural change), takes to settled agriculture, and clears the land for this purpose (an environmental change), the following changes may follow. Agricultural activity and a settled population may offer an excellent breeding ground for malarial mosquitoes. Therefore, incidents of malaria will increase, as will the proportion of individuals suffering from sickle cell anemia. The latter individuals have certain physical disadvantages but are immune from malaria, and their immunity is genetically transmissible (a genetic change). Thus, culture, environment, and the genetic character of the population change together.

Although the vectors just described cannot be isolated in practice, they are heuristically desirable because they introduce a degree of tidiness into psychologists' thinking. Such fine distinctions seldom intrude into psychologists' experimental work. Most studies of perception are not concerned with the genetic vector at all; indeed, only a few attempts to separate the effects of the environment from the effects of culture have been made. This is not to say that results of such studies lack merit but rather to stress that they provide rather confounded data so far as the three vectors are concerned. However, such theoretically impure studies are of immediate practical use because they offer guidance, for example, for selection of personnel when the causes of particular abilities are of no consequence but the presence or absence of such abilities is.

3. VISUAL ILLUSIONS

One of the phenomena investigated by Rivers in the course of the Torres Straits expedition was that of perception of visual illusions. It was a study intended to compare "Western" and "non-Western" populations on a number of illusion figures that, at the time, were thought to constitute a homogeneous category of stimuli. It was not driven by any clearly defined hypothesis. The same phenomena were investigated by Segall and colleagues more than 60 years later. However, their investigation was inspired by specific hypotheses concerning the extent to which environmental experience affects perception. Segall and colleagues sampled a number of distinct populations inhabiting distinct environments and having different cultural artifacts. The conclusions of the two studies were in agreement in that both suggested that common visual illusions do not form a homogeneous body of phenomena and that their effects differ among populations.

Segall and colleagues attributed the differences observed to the differences in visual experience, for example, to openness of vistas (i.e., the extent to which people can see open spaces—savannah dwellers can do so to a larger extent than can city dwellers) in the case of the horizontal–vertical illusion and to "carpentered environment" (i.e., the extent to which artifacts of a culture incorporate straight edges—Western artifacts do so to a larger extent than do those of, say, Zulus) in the case of the Muller–Lyer illusion. It should be noted that the cultural vector is explicitly involved in the latter case. Of course, it is not entirely absent in the former case.

4. PICTORIAL PERCEPTION

Surprising responses of individuals drawn from exotic populations to pictures have frequently been mentioned by travelers and anthropologists. These have ranged from reports of the complete inability of certain populations to perceive that pictures represented objects to reports of responses indicating that the depicted objects were seen as real (e.g., that the portrayed people were alive but asleep). These observations, which came from reputable observers, were casual and not based on data obtained in the course of properly conducted studies.

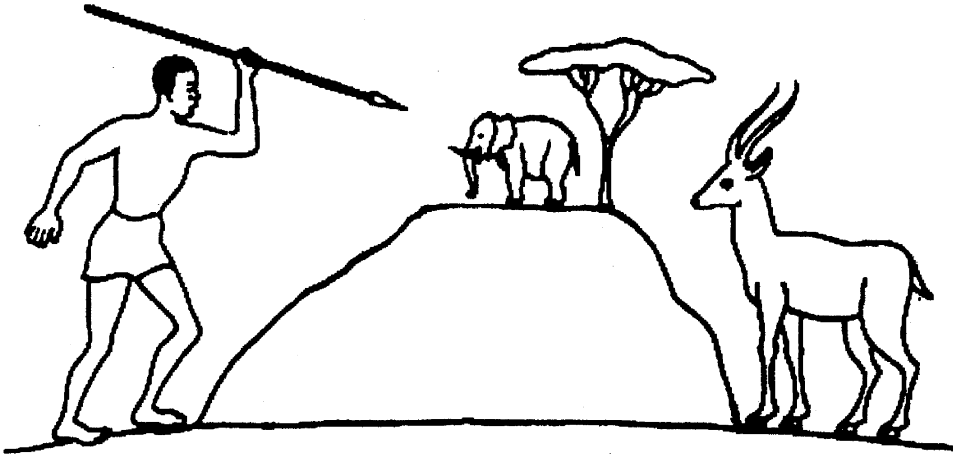


FIGURE 1 One of Hudson's items: A test of pictorial perception.

The problem was first tackled systematically when misperception of pictures was found to be an obstacle to working in industry. It was found that the difficulties experienced were those associated with the failure to perceive the third dimension in portrayals. This was detected by asking individuals to name distinct elements in the test picture, such as the man, the spear, and the duiker in Fig. 1 (taken from Hudson's test), and, if they named them correctly, asking for simple judgments of depicted depth. The questions put to the individuals were "Which animal is nearer to the hunter?" and "At which animal is the hunter aiming his spear?" The answer "elephant" to either of these questions was taken as showing difficulties in perception of pictorial depth.

Following this study, several other measures were developed. These measures did not rely on verbal responses but instead called for reconstructing the pictured scenes or for copying an "impossible figure," taking advantage of the fact that the impossible figure in question is seen as "impossible" only if it is seen as a depiction of a three-dimensional object.

Because pictures are such an important means of communication whenever linguistic communication is not possible, and because the population in such cases is illiterate or has a language that is not widely known, these findings aroused widespread interest. A number of investigations were carried out, intent on defining the phenomenon, identifying its causes, and determining the most efficacious portrayals. It was found that perception of pictorial depth is not an "either/or" phenomenon but rather is greatly affected by the nature of pictorial stimuli, so that certain geometrical figures (notably certain illusion figures) evoke perception of depth even though they are not reminiscent of any objects and are not familiar to the observers. In contrast, certain other figures (e.g., silhouettes) may readily evoke recognition without evoking perception of depth. An example of the latter is the elephant in Fig. 1, which is recognized as an elephant (a three-dimensional entity) but is not seen as three-dimensional. Indeed, even so-called "pin men" or "stick men," consisting merely of a set of straight lines that depict arms, legs, and trunks, are readily recognized as depictions of men, although they contain no depth cues whatsoever. The term "epitemic" was proposed as a descriptor of such figures that portray objects but lack perceptual depth, whereas the term "eidolic" was proposed for those depictions that evoke perception of depth, although they do not portray anything recognizable (e.g., some illusions).

It was also found that formal schooling is associated with the ability to perceive pictorial depth. Therefore, it has been put forward that pictorial perception is a matter of skill. However, it was also found that otherwise very capable engineering students drawn from certain populations were markedly poor at engineering drawings and that unschooled illiterate Bushmen's performance on the Kwengo calipers test (in which individuals must set a pair of calipers to a depicted angle) varied with ethnic origin and, in the case of one Bushmen population, did not differ from the performance of urbanized and regularly schooled Zulus.

It is noteworthy that intercultural differences were reported on other visual tasks. Differences were obtained on measures of implicit shape constancy (i.e., shape constancy as perceived in pictures). African

individuals were less affected by this constancy than were European individuals. Kwengo Bushmen's children, Zulu children, and White South African children were found to differ greatly in their responses to incomplete figures. The Kwengo performed better than did the two other samples. It is highly probable that they did so because they were better at integrating the stimuli. However, the origin of this skill (if in fact it is a skill) is not apparent; Bentley, influenced by Berry, thought that it might be a facet of the skill of hunting. Therefore, it seems plausible that the difficulties of pictorial perception can arise from failure to integrate perceptual input in the manner intended by the artist. The ease with which young children copy "impossible" figures that contain mutually contradictory elements, and the fact that those individuals from relatively pictureless cultures who build well-integrated three-dimensional models find it difficult to copy drawings containing mutually contradictory cues, supports this view.

The reported investigations skipped over the most striking of pictorial failures, that is, the failure to recognize that a picture portrays an object. There are several anecdotal reports of this but only a paucity of systematic evidence. The evidence in hand suggests that such difficulties are rare and further suggests that an inexperienced pictorial perceiver may attach inordinate weight to the material used in making the picture. For example, if paper is not a common cultural artifact, the paper on which a picture is made, rather than the picture itself, may grasp the observer's attention. Therefore, the picture recognition difficulties may be reduced by substituting a familiar material for an unfamiliar one (e.g., substituting coarse cloth for paper).

The pragmatic implications of these findings are clearly those known to every advertiser: It is imperative that pictures be carefully evaluated on a sample of the population for which they are intended.

5. PERCEPTION OF PATTERNS

Another aspect of cross-cultural perception that attracted attention is that of reproduction of geometric patterns. It came to the fore when psychological tests such as Kohs blocks were used and scores attained by some populations were found to be so low that the effectiveness of the tests became questionable. In some instances, the tasks were modified to strengthen their interindividual discriminability. In Papua New Guinea, cubes used to reproduce patterns were replaced by flat plaques and trays that included appropriate recesses in which the patterns were to be reproduced. In South Africa, a newly designed and interesting test required individuals to render symmetrical, by placement of a rectangular fiche, an arrangement of three identical rectangles. Both of these approaches were used for selection, the former in a scholastic setting and the second in an industrial setting.

Suppression of a recalcitrant variable, although justified by practical considerations, does not elucidate the theoretical problem. A number of investigations concerned factors such as the symmetry of the pattern and its apparent stability as well as individuals' tendency to leave the pattern in whatever orientation it was completed. A persistent reported tendency was to rotate the reproductions of patterns in such a way as to increase both their symmetry about the observers' median planes and their apparent stability. The question of whether the errors of orientation are errors of inattention to orientation (where the observers left the correctly reproduced pattern in whatever orientation they happened to complete it) or result from some lawful perceptual vector was addressed by Biesheuvel and Jahoda.

There are cultural groups, such as the Guugu Yimithirr of Northern Queensland, who judge orientations by referring to the geographical coordinates rather than to those of their own bodies (e.g., left and right). The implications of the two types of judgment are demonstrated in Fig. 2, which shows a stimulus viewed from the distinct stances. Those relying on the geographical notions see the orientation of the stimulus as the same from both stances (it is consistently pointing westward). Those relying on their body coordinates see the orientations as different; from Stance A, it points to the left, and from Stance B, it points to the

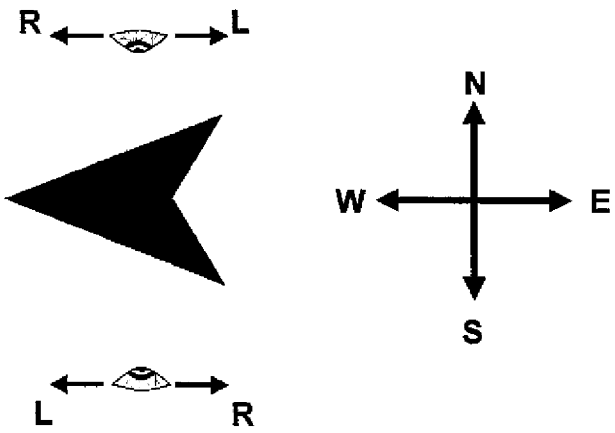


FIGURE 2 Graphical explanation of the effects of the two distinct frameworks on the object's reported orientation.

right. Two issues that ought to be noted here are (a) the justification of describing the phenomenon as perceptual and (b) that arising from the psycholinguistic gloss that would seek the roots of the phenomenon in linguistic use (i.e., "it is consistently pointing westward" vs "it is consistently pointing to the left").

6. EIDETIC IMAGERY

Some people, notably children, experience a special kind of "photographic" imagery, pictorial in nature, that enables them to report with striking accuracy. This eidetic imagery was once believed to be more widespread in illiterate populations. However, empirical data do not support such a notion, and the very phenomenon remains one of the puzzles of psychology.

7. STUDIES OF PERCEPTION OF COLOR

Cross-cultural studies of perception of color have been dominated from their very beginning by the issue of the relationship of color vocabularies to perception of color. If the color spectrum can be conceived as a continuum resulting from steadily changing frequency of light waves, such a continuum can be arbitrarily divided in any manner and different languages may present different ways in which to divide it. Speakers of languages dividing the continuum differently will respond to colors differently. For example, speakers of a language that uses the same term for blue and green will respond to these two colors identically. On the other hand, if there is no such continuum, certain frequencies may be perceived as belonging together. That is, if observers, independent of the language, perceive pale green and dark green as mutually related and perhaps perceive a certain kind of green as the epitome of green (called focal green), the color spectrum obviously is not perceived as a uniformly changing continuum and perceived colors themselves impose certain differences in nomenclature. Neither of these polar positions has been sustained empirically, and the fact that pure colors can be blended to evoke percepts identical to those evoked by other pure colors erodes the notion of a perceptual continuum similar to that of, say, weight.

There is empirical evidence that linguistic groups tend to perceive colors differently in accordance with the categorization prevalent in their languages. For example, some Bantu languages use the same word for green and blue, and speakers of these languages tend to treat these colors as if they were one; however, these effects are not as strong as rigorous application of the linguistic hypothesis would lead one to expect. There is also evidence that when focal colors are sought, these are identical in whatever language is used by the individuals. An additional complication arises because the same color may have quite different significances depending on the object of which it is a color. For example, the focal color detected when using Munsell chips may be quite different from that detected when using toy buffaloes. This effect is likely to be influenced by cultural differences in attention paid to color as opposed to shape of stimuli.

The notion that cross-cultural differences may arise due to differences in the structure of the eyes, namely in the pigmentation of the fundus oculi, has been put forward and investigated and has not yielded convincing data.

8. PERCEPTION OF TIME

Travelers' tales about surprising notions of time and of timekeeping are about as numerous as those about perception of pictures, yet the former have not attracted a comparable number of investigators, perhaps because the concepts involved are far more complex. Not only can a distinction be made between "social" time (e.g., as measured by punctiliousness, punctuality, or adherence to time schedules) and "physiological" time (e.g., as measured by the ability to judge and reproduce time intervals), but it also appears that attitude toward time and time-related concepts may vary between urban and rural groups and may affect the memorability of such concepts.

9. CONCLUDING REMARKS

It is apparent from this article that the dominant characteristic of cross-cultural studies in perception is their unevenness. Certain areas are pursued with vigor, whereas others are entirely neglected. For example, there are no extensive studies of basic perceptual processes. This is unfortunate because there are issues in relation to which only comparisons of distinct groups of individuals can define the extent of human capacities. For example, the surprising performance of the Bushmen on spatial tasks calls for a more fundamental explanation than does a statement that it may be associated with their hunting practices. These considerations extend beyond the here and now because

cross-cultural studies of perception also embrace studies of cultural artifacts of populations long vanished and sometimes known only by their artifacts. Such are the studies of Paleolithic art, that is, the studies of the early cognitive acts of humans.

It is noteworthy that none of the cultural differences demonstrated suggests that a population has unique perceptual attributes. The differences are in the intensity or frequency of attributes. When reflecting on such differences, one must bear in mind individual differences that are inevitably present in every population.

Probably the most important contribution that cross-cultural studies of perception can make is that of determining the range of human perceptual functions. Determination of the true range is essential both for a full understanding of such functions and for establishing a universal taxonomy of such functions.

See Also the Following Articles

Conformity across Cultures ■ Cross-Cultural Psychology, Overview ■ Industrial/Organizational Psychology across Cultures

Further Reading

Deręgowski, J. B. (1980). *Illusions, patterns, and pictures: A cross-cultural perspective*. London: Academic Press.

Deręgowski, J. B. (1989). Real space and represented space: Cross-cultural perspectives. *Behavioral and Brain Sciences*, *12*, 51–119.

Levinson, C. S. (1998). Studying spatial conceptualization across cultures: Anthropology and cognitive science. *Ethos*, *26*, 7–24.

Reuning, H., & Wortley, W. (1973). Psychological studies of the Bushmen. *Psychologia Africana* (Monograph Suppl. No. 7).

Rivers, W. H. R. (1901). Vision. In A. C. Haddon (Ed.), *Reports of the Cambridge anthropological expedition to Torres Straits* (pp. 1–132). Cambridge, UK: Cambridge University Press.

Russell, P. A., Deręgowski, J. B., & Kinnear, P. R. (1997). Perception and aesthetics. In J. W. Berry, P. R. Dasen, & T. S. Saraswathi (Eds.), *Handbook of cross-cultural psychology* Vol. 2: *Basic processes* (pp. 107–142). Boston: Allyn & Bacon.

Segall, M. H., Campbell, D. T., & Herskovits, M. J. (1966). *The influence of culture on visual perception*. Indianapolis, IN: Bobbs–Merrill.

Performance Slumps in Sport: Prevention and Coping

J. Robert Grove
University of Western Australia, Perth, Australia

1. Slumps Versus Choking and Plateaus
2. Causes of Performance Slumps
3. Psychological Consequences of Slumps
4. Factors Influencing the Reaction to Slumps
5. Coping Behaviors in Response to Slumps
6. Can Slumps Be Prevented?
7. Increasing Awareness and Managing Risk
8. Maintaining Confidence and Emotional Stability
9. Performance Recovery Strategies
10. Summary and Conclusions
Further Reading

GLOSSARY

attribution The perceived cause(s) of an event; these perceived causes have both emotional and motivational significance, so they can influence an athlete's behavior in response to performance difficulties.

choking Temporary unexpected decrements in performance that occur in situations where the perceived importance of good performance is high.

coping responses The specific cognitive and behavioral strategies that an individual uses to manage distress; these strategies are sometimes categorized in terms of their general approach or avoidance qualities.

motivational climate The psychosocial environment surrounding training and competition; rivalry, punishment, and recognition based on superiority to others are generally viewed as elements of a counterproductive climate, whereas cooperative interaction and recognition based on improvement are generally viewed as elements of a productive climate.

performance profiling Systematic analysis of the essential ingredients for effective performance in a particular setting in combination with an assessment of strengths and weakness along these lines; this profiling can include physical, mental, and/or strategic aspects of performance and is useful for prioritizing skill development needs.

plateau Stabilization of performance at a given level, that is, a failure to improve.

process goals Specific and measurable objectives that relate to what one will do and how one will do it rather than the desired result or outcome; shooting 100 free throws at the end of practice every day for a month is an example of a process goal related to basketball.

relapse prevention A systematic attempt to reduce the probability that one occurrence of an undesirable behavior will lead to long-term debilitation or repeated occurrences; this framework emphasizes awareness of high-risk situations as well as the development of cognitive–behavioral strategies to maintain feelings of self-efficacy and control.

self-efficacy Situation-specific confidence in one's ability to accomplish something.

self-handicapping Generation of excuses or potential obstacles in advance of performance; although such actions may reduce the emotional impact of poor performances, they can also set in motion a self-fulfilling prophecy that makes good performance less likely.

©2004 Elsevier Inc.
All rights reserved.

Slumps are extended periods of below-average performance. The usual level of one's performance provides the frame of reference for defining a slump rather than some idealized or desired level of performance. The nature of one's sport involvement is also an important consideration. For a high-level athlete who trains daily and competes regularly, 2 or 3 weeks of below-average performance could constitute a slump. For a less committed performer with an irregular practice and competition schedule, a similar period of poor performance might not constitute a slump.

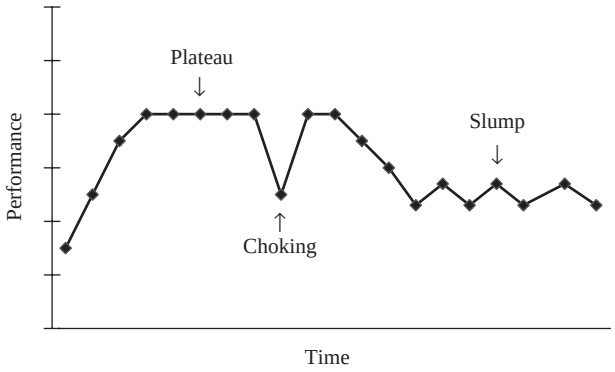


FIGURE 1 Illustration of the distinction among plateaus, choking, and slumps.

1. SLUMPS VERSUS CHOKING AND PLATEAUS

Individuals who participate in sport devote considerable time and effort to the development of physical skills, and they derive personal satisfaction from the public display of these skills. Consistency of physical performance is especially important for high-level amateur and professional athletes. Their scholarships or contracts may include performance-related contingency clauses, and their self-identity is often closely linked to physical prowess. Therefore, substandard performances are a source of concern for these individuals, and prolonged periods of poor performance can be highly stressful. Recreational athletes may also experience threats to their sense of self when, for no obvious reason, customary levels of performance are not attained for an extended period of time. The term "slump" is used to describe such a state of affairs.

Slumps can be distinguished from the related but qualitatively distinct phenomena of choking and plateaus. Choking refers to temporary unexpected decrements in performance that occur in situations where the perceived importance of good performance is high. Although choking can occur repeatedly, it is distinguishable from a performance slump because performance levels return to normal when the perceived importance of the event decreases. Factors associated with perceived importance include performance-contingent rewards or punishments and the ego relevance of the event. Performance plateaus are a longer term phenomenon than is choking, and the former are characterized by a failure to improve rather than by a decrement in performance. The athlete's performance improves to a point but then, despite diligent and conscientious training efforts, remains at a given level for an extended period of time. The distinction among plateaus, choking, and slumps is illustrated in Fig. 1.

2. CAUSES OF PERFORMANCE SLUMPS

The inability to pinpoint a specific cause for the extended period of poor performance is an important phenomenological aspect of performance slumps. Frequently (but not always), there is no apparent explanation for the performance deterioration, and this ambiguity creates frustration and a sense of helplessness for the athlete, who feels like he or she is working harder and harder but, despite this additional effort, is making little or no progress. However, from both theoretical and experiential perspectives, performance slumps are typically believed to originate from physical, psychological, behavioral, technical, and/or environmental causes. Table I summarizes these general categories of causal factors and provides examples of each of them. It is important to recognize that these categories are not mutually exclusive. In other words, performance slumps can have multiple causes, and it is entirely possible for factors within one category to have an influence on factors in another category.

Physical causes include lack of fitness, illness, injury, and/or general fatigue. With respect to fitness, it is important to recognize the distinctive purposes of general physiological fitness (e.g., strength, aerobic capacity) and of competition-specific fitness. Athletes sometimes have a tendency to place a heavy emphasis on general fitness and overlook the importance of match-specific fitness. This tendency is especially

TABLE I
Possible Causes of Performance Slumps

Category	*Subcategories*	*Examples*
Physical factors	Fitness	Training not specific to game conditions
	Illness or injury	Lingering cold or flu symptoms
	Fatigue	Suboptimal work-to-rest ratio
Psychological factors	Confidence	Uncertainty about style of play or role
	Concentration	Poor control of distractions and/or emotions
	Expectancies	Unrealistic demands from self/others
Behavioral factors	Routines	Inconsistent preperformance activities
	Self-handicapping	Precompetition behaviors that create obstacles
	Time management	Overcommitment and/or poor organization
Technical factors	Technique/Equipment	Minor changes that disrupt "system integrity"
	Coaching	Too much and/or confusing instruction
	Knowledge	Poor awareness of on-field contingencies
Environmental factors	Motivational climate	Excessive focus on outcomes or superiority
	Facilities	Substandard training/playing conditions
	Nonsport stressors	Family, work, or school problems

strong early in the competitive season when there is a need to make a transition from general off-season programs to performance-specific programs. Fatigue and illness/injury, on the other hand, can become problematic as the season unfolds. The cultural milieu within competitive sport is typically one of "hard work pays dividends," so athletes often push themselves to the brink of exhaustion in an effort to improve performance. A failure to recognize that optimal performance requires a balance between work and rest can produce decrements in performance that persist for an extended period of time.

From a psychological perspective, the ability to play with absorption and certainty of purpose is often viewed as the cornerstone of performance excellence. If athletes do not have a clear idea of their on-field or on-court roles, their performances can suffer due to hesitancy ("lack of conviction"). The heavy emphasis on results in competitive sport also makes it easy for athletes to lose sight of their performance objectives and become overly concerned with outcomes. This loss of focus can be both subtle and debilitating, and it is especially likely to occur when significant others (e.g., coaches, teammates, family) emphasize achievements when interacting with the athletes. Persistent decrements in performance can also occur if athletes allow themselves to become distracted by emotional overreactions and/or the behavior of teammates, opponents, fans, or officials. Tendencies along these lines are sometimes rationalized as "competitive intensity" and so can persist despite being counterproductive.

Consistent performance in stressful competitive circumstances requires a consistent approach to training and performing. Schedules, plans, and routines help to provide this consistency, and performance disruptions can occur if these elements are missing, underdeveloped, or changed. Seemingly innocuous changes, such as alterations in training times, can sometimes interfere with performance by producing a cascade effect whereby many other activities and commitments must be rearranged to fit the new schedules. Similarly, inadvertent departures from standard precompetition or preexecution routines can sometimes lead to persistent performance decrements. Self-handicapping behaviors can also interfere with performance and can infiltrate the athlete's preparatory activities in a subtle and seemingly defensible manner. Common self-handicapping acts include paying increased attention to sport-specific issues that are unrelated to performance, training excessively hard close in time to competition, and being preoccupied with minor injuries or illnesses. Failure to manage one's time effectively can also undermine performance by creating a general sense of urgency about whatever the athlete is doing at the time.

Technical factors that can contribute to performance slumps include changes to technique and/or equipment. Although sometimes such changes are not obvious, they can disrupt the integrity of the performance

system by altering finely tuned timing mechanisms. Attempts to execute skills in an ideal or prescribed manner can have a similar disruptive influence on performance that might take considerable time to overcome. Coaches sometimes contribute to this problem by offering too much input, providing overly technical input, and/or by showing a lack of flexibility with respect to methods. A lack of awareness regarding when and how to make on-field adjustments to strategy can also contribute to extended periods of subpar performance despite the presence of objectively good technique. Potential causes of slumps within the environment include substandard training or competition facilities, stressors related to nonsport roles, and a motivational climate that places a heavy emphasis on beating others and winning rather than on skill development.

3. PSYCHOLOGICAL CONSEQUENCES OF SLUMPS

At some level, athletes acknowledge performance slumps as an "occupational hazard." At the same time, they admit to experiencing considerable distress when confronted with them. Prominent psychological reactions to slumps include feelings of frustration, anger, helplessness, self-doubt, and anxiety. Frustration is known to increase when there is proximity to a desired goal but attainment of the goal is blocked. For slumping athletes, the fact that they are accustomed to performing at a higher level serves as a continual reminder that the goal can be attained. On the other hand, their inability to reestablish these levels of performance despite diligent effort is a source of considerable frustration. This frustration is accompanied by increased emotionality, which may manifest itself in the form of impatience, irritability, and/or hostility. It may also manifest itself in feelings of helplessness, despair, and depression.

Loss of confidence, self-doubt, and anxiety are also experienced during a performance slump. It has been shown repeatedly that the strongest influence on situation-specific confidence (i.e., self-efficacy) is prior performance in similar situations. When the most recent prior performances have been consistently poor, confidence is eroded. At the same time, physical prowess is an important component of self-identity for most athletes, and there is no shortage of physically talented individuals waiting in the wings. Performing at less-than-expected levels for an extended period can threaten the athlete's sense of self and lead to pessimistic rumination on things such as "Do I still have what it takes?" and "Will I lose my spot [on the team, in the starting lineup, in the tournament, etc.]?" Such thoughts reflect self-doubt as well as anxiety about the future.

4. FACTORS INFLUENCING THE REACTION TO SLUMPS

Research findings indicate that the amount of stress experienced in response to a slump is related more to how the slump is appraised than to individual difference factors. More specifically, self-reports from athletes in a variety of sports indicate that slump-related stress is slightly lower for those with spouses or partners than for those without spouses or partners but that it is unrelated to age, sex, education level, or years of competitive experience. On the other hand, interpreting the slump as a loss or a threat substantially increases the level of perceived stress. Slumps are also perceived as stressful when they are interpreted as a challenge, but not to the same degree as when loss or threat interpretations are involved. Dimensional properties of causal interpretations also show some relationship to slump-related stress. Specifically, a lack of perceived control over the cause of a slump tends to increase the level of self-reported stress, whereas a high level of perceived control over the cause tends to decrease stress levels. This finding is, of course, consistent with the way in which perceptions of control are viewed within general theories of behavior change as well as within models of stress and coping.

5. COPING BEHAVIORS IN RESPONSE TO SLUMPS

Athletes tend to be concrete thinkers, and the sport environment is one that reinforces hard work and persistence. Therefore, it is not surprising to find that active, problem-focused coping strategies are cited most often when athletes are asked about their behavior in response to a slump. More specifically, findings from both qualitative and quantitative studies using various coping inventories indicate that athletes frequently employ increased effort or resolve, planned action, seeking of technical advice, seeking of social

support, and suppression of competing activities when they are confronted with an extended period of poor performance. Although emotion-focused strategies are used less often, they are sometimes employed. When that is the case, strategies related to conscious control of emotions, self-blame, and wishful thinking appear to be used most often. Avoidance-oriented strategies, such as denial and behavioral disengagement, are consistently the least frequently cited coping behaviors by athletes confronted with performance slumps.

Even though athletes have a general tendency to cope with performance slumps as outlined above, gender and personality influences are also evident in the literature. With respect to gender, it has been shown that female athletes seek emotional social support to a greater extent than do male athletes in response to a slump, but these differences tend to disappear when instrumental social support is examined (or when the two types of social support are considered together). Other studies have shown that male athletes report more use of the avoidance-oriented strategies of humor, denial, and behavioral disengagement than do female athletes. It must be emphasized that the overall frequency of citation for these strategies is low among athletes, but the findings are interesting because they run counter to findings outside of sport that show a preference for active, approach-oriented coping strategies among males.

Personality factors that have been examined in connection with slump-related coping include dispositional optimism, trait sport confidence, perfectionism, self-consciousness, and achievement goal orientation. Table II summarizes the relationships that have emerged in this research. In general, athletes who are optimistic, possess high levels of domain-specific confidence, and adopt a mastery orientation to achievement tend to use active, problem-focused strategies in an effort to overcome performance difficulties. At the same time, these athletes are disinclined to rely on avoidance-based strategies, such as wishful thinking and behavioral disengagement, in response to a slump. Highly self-conscious athletes, on the other hand, have a tendency to fantasize, become preoccupied with their emotions, blame themselves for their difficulties, and/or disengage from the task. The impact of perfectionism on slump-related coping appears to depend on the particular subcomponent that predominates. More specifically, when perfectionist tendencies are driven by excessive concern about mistakes,

TABLE II
Personality Correlates of Slump-Related Coping among Athletes

Personality construct	*Coping strategies*	*Nature of relationship*
Dispositional optimism	Effort/Resolve	Positive
	Planned action	Positive
	Seeking social support	Positive
	Wishful thinking	Negative
	Behavioral disengagement	Negative
Trait sport confidence	Deliberate emotional control	Positive
	Behavioral disengagement	Negative
Mastery achievement goals	Planned action	Positive
	Behavioral disengagement	Negative
Self-consciousness	Wishful thinking	Positive
	Self-blame	Positive
	Emotional focus/venting	Positive
	Behavioral disengagement	Positive
Perfectionism		
Concern over mistakes	Denial	Positive
	Behavioral disengagement	Positive
High personal standards	Suppress competing activities	Positive
	Planned action	Positive
	Self-blame	Positive
	Emotional focus/venting	Positive

avoidance-based coping strategies are employed. On the other hand, when concern over mistakes is accompanied by a focus on high personal standards, a potentially more productive mix of problem-focused and emotion-focused strategies tends to be used.

6. CAN SLUMPS BE PREVENTED?

In general, the answer to this question is no. Individuals who participate in competitive sport on a regular basis will, in all likelihood, be confronted by extended periods of below-average performance from time to time. Acknowledging this possibility from the outset has at least two beneficial consequences. First, it reinforces the need for athletes to understand why slumps occur and to approach their sport involvement in ways that will minimize the frequency of the occurrence of slumps. Second, it suggests that athletes also need to learn about the typical reactions that occur during a slump and need to develop behavioral strategies for "getting back on track" as quickly as possible. To put it somewhat more bluntly, athletes are probably going to be confronted with performance slumps from time to time, so they had better learn how to deal with them productively. In the area of substance addiction, this is the same philosophy that underlies relapse prevention training, and there is merit in using a similar approach in connection with performance slumps. Figure 2 outlines a proposed structure for using the relapse prevention model in this way. Specific strategies for increasing awareness, maintaining confidence and emotional control, and facilitating a return to form are discussed next.

7. INCREASING AWARENESS AND MANAGING RISK

The first step toward increased awareness and effective risk management is to become knowledgeable about behaviors and circumstances that increase the potential for slumps to occur. The list of factors in Table I, although not exhaustive, is helpful in this regard because it highlights some of the more common features of high-risk situations. These "red flags" include depleted physical reserves, information overload, inconsistent training and competition routines, a chaotic and disorganized lifestyle, and the presence of unusually strong nonsport stressors related to work, family, or educational duties. Specific interventions that address these issues are outlined in Fig. 3 and include self-monitoring of physical and mental state, systematic observation of coach–athlete interactions, establishment of regular daily and weekly schedules, rehearsal of preperformance routines, and development of time management skills.

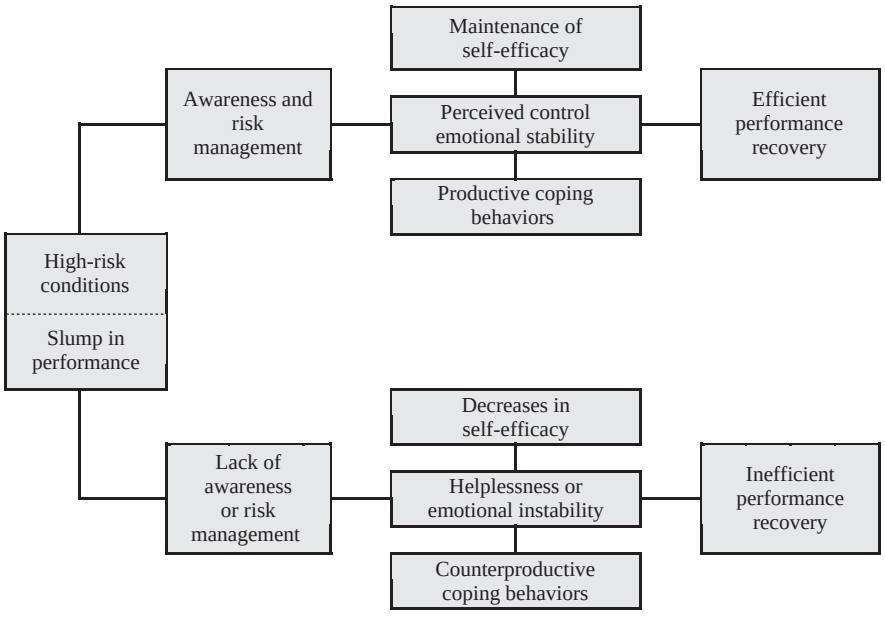


FIGURE 2 Application of the relapse prevention model to performance slumps in sport.

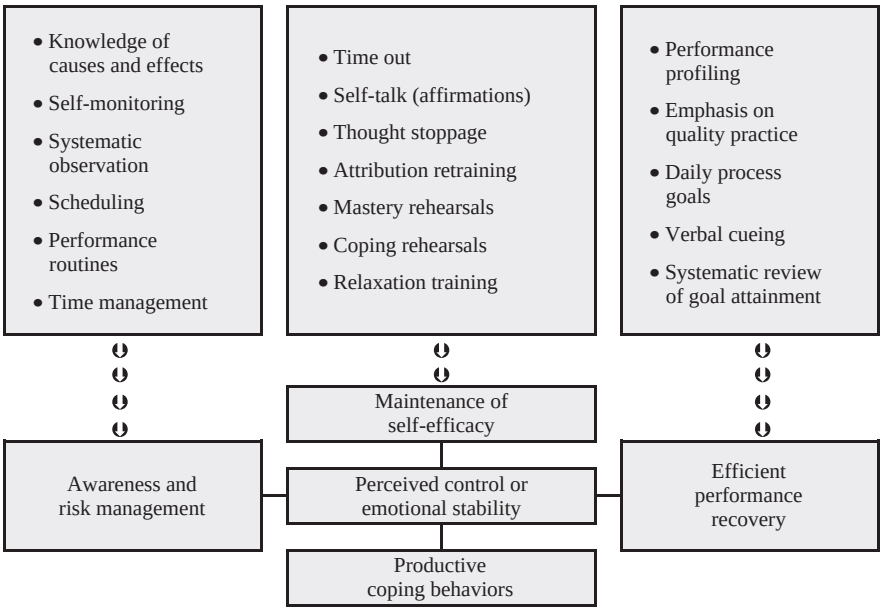


FIGURE 3 Intervention strategies for the management of performance slumps.

8. MAINTAINING CONFIDENCE AND EMOTIONAL STABILITY

When athletes are performing poorly, it is natural for them to lose confidence, feel somewhat helpless, and react with heightened emotionality. If not controlled, these negative thought patterns and emotional reactions can lead to a self-perpetuating cycle of counterproductive cognitive–behavioral reactions. To gain control over these reactions, it is sometimes necessary for athletes to remove themselves from the competitive environment for a period of time. This "time out" period, which might last from a few days to a few weeks depending on the level of involvement, gives them a chance to "clear their heads" and stabilize their emotions. There is certainly value in doing so, but it is sometimes difficult to convince committed sport performers that the best thing they can do for their game is to forget about sport for a while. An examination of Table II suggests that this might be especially problematic for highly optimistic/confident athletes with a strong mastery orientation.

Selected thought management strategies can be implemented to maintain feelings of self-efficacy and control as well as emotional stability. These strategies include affirmative self-talk, thought stoppage, and cognitive restructuring of attributions for the slump. Affirmations are simple, declarative, first-person statements that reinforce positive personal qualities. As such, they are antithetical to the self-doubts that typically accompany performance difficulties. Examples of affirmations that might be relevant during a slump include "I deserve to be here," "I've worked hard at my game," "I can produce when I have to," and "I enjoy showing what I can do." A simple, three-part thought stoppage procedure can be used either by itself or in combination with affirmative self-talk to block negative thoughts and redirect attention in a constructive manner. This procedure consists of using covert verbal triggers (e.g., "Stop") and/or physical triggers (e.g., snapping one's fingers) to arrest negative thoughts. Attention is then redirected through a brief focus on the physical self (e.g., "Breathe . . .") and/or the use of a cue word or phrase that captures a desired performance quality or orientation (e.g., "Let it rip," "Play your game"). Affirmations can, of course, be used as the cue phrases in this procedure.

Attributions play an important role in the way in which athletes respond to performance difficulties both emotionally and motivationally. In general, it is emotionally comforting and motivationally productive to attribute poor performance to unstable specific factors (e.g., an unacceptably poor first serve percentage in tennis) rather than to stable global factors (e.g., a general inability to serve well). Doing so reduces the emotional impact of the negative outcomes and sets the

stage for continued striving toward improvement of the specific weakness. These explanations must be realistic, of course, because contrived excuse making will certainly not be advantageous. At the same time, consultants should be aware of the link between attributions and feelings of helplessness, and they should remain vigilant for potentially counterproductive attributional patterns when persistent performance difficulties arise.

Mental rehearsal and a variety of relaxation procedures may also be beneficial for maintaining confidence and emotional control. Two different types of mental rehearsal are relevant and can be used to one's advantage at different times. Mastery rehearsal involves visualization of effective performance and is good for maintaining self-confidence. It can be implemented by asking athletes to recall an excellent performance from the past and to reexperience it in as much detail as possible. After doing so, they can be asked to visualize an upcoming performance situation, see themselves performing in a similarly effective manner, and associate the prior feelings of success with the future performance. Coping rehearsal, on the other hand, involves visualization of potential difficulties that might be encountered in the future and the strategies that could be used to manage these difficulties effectively. This form of imagery aids in the recognition of behavioral alternatives and in the development of response flexibility. Relaxation exercises of various kinds can be used in isolation or in combination with mental rehearsal to alleviate some of the tension that typically accompanies a series of frustrating performances. Abdominal breathing exercises and abbreviated forms of progressive muscular relaxation are good choices for athletes because they focus attention on physical processes to produce a relaxation response. Sport performers are physically oriented people, and they are sometimes more accepting of physically oriented relaxation techniques than of cognitively oriented techniques.

9. PERFORMANCE RECOVERY STRATEGIES

On-field performance is the bottom line for sport performers, so taking steps to increase the probability of efficient performance recovery is an important element of slump-related coping. A variety of behavioral techniques are relevant here. For example, performance profiling can be used to isolate key elements of the physical and/or mental game as well as to identify strengths and weaknesses in these key areas. Skill development prioritization follows quite naturally from such an exercise, so the stage is set for quality practice (as opposed to "mindless" practice). Such an approach to training has been documented as a core element of sporting excellence, and there is considerable merit in modeling performance recovery strategies on these same principles. A central feature of quality practice is the incorporation of daily process goals into the training regime so that athletes have a clear idea of what they will do on any given day, how long they will do it, and how they will evaluate what they have done. Verbal cueing can be used along with daily goal setting to remind athletes of their priorities and to direct attention toward processes rather than outcomes. Cueing can be done covertly by athletes through self-talk, overtly by athletes through written reminders, or overtly by informed others through verbal statements. Structured reviews of the extent to which process goals were attained can then take place following training sessions and/or competitive performances.

10. SUMMARY AND CONCLUSIONS

Slumps are an inherent and unavoidable phenomenon within physical performance settings. They are qualitatively distinct from the related phenomena of choking and plateaus, and they have predictable psychological consequences. These consequences include decreases in self-efficacy, feelings of helplessness and loss of control, and behavioral attempts to cope through approach or avoidance. Individual difference factors that have been shown to influence slump-related coping among athletes include gender and personality. Because it is likely that most sport performers will experience slumps at some points in time, it would be prudent for them to adopt a risk management perspective on the process. Such a perspective involves (a) increasing awareness of the potential causes of and common reactions to slumps, (b) becoming familiar with cognitive–behavioral strategies for the maintenance of confidence and emotional stability in the face of performance difficulties, and (c) adopting a process-focused approach to training and competition. Doing so will help to reduce the frequency of occurrence of slumps and will promote efficient restoration of typical performance levels when they do occur.

See Also the Following Articles

Coping ■ Intrinsic and Extrinsic Motivation in Sport ■ Self-Confidence in Athletes

Further Reading

Baumeister, R. F., & Showers, C. J. (1986). A review of paradoxical performance effects: Choking under pressure in sports and mental tests. *European Journal of Social Psychology*, *16*, 361–383.

Beilock, S. L., & Carr, T. H. (2001). On the fragility of skilled performance: What governs choking under pressure? *Journal of Experimental Psychology: General*, *130*, 701–725.

Conroy, D. E., & Benjamin, L. S. (2001). Psychodynamics in sport performance enhancement consultation: Application of an interpersonal theory. *The Sport Psychologist*, *15*, 103–117.

Eklund, R. C., Grove, J. R., & Heard, N. P. (1998). The measurement of slump-related coping: Factorial validity of the COPE and modified COPE inventories. *Journal of Sport & Exercise Psychology*, *20*, 157–175.

Goldberg, A. S. (1998). *Sports slump busting.* Champaign, IL: Human Kinetics.

Grove, J. R., Eklund, R. C., & Heard, N. P. (1997). Coping with performance slumps: Factor analysis of the Ways of Coping in Sport Scale. *Australian Journal of Science and Medicine in Sport*, *29*, 99–105.

Grove, J. R., & Heard, N. P. (1997). Optimism and sport confidence as correlates of slump-related coping among athletes. *The Sport Psychologist*, *11*, 400–410.

Grove, J. R., & Stoll, O. (1999). Performance slumps in sport: Personal resources and perceived stress. *Journal of Personal and Interpersonal Loss*, *4*, 203–214.

Klinger, E., Barta, S. G., & Glas, R. A. (1981). Thought content and gap time in basketball. *Cognitive Therapy & Research*, *5*, 109–114.

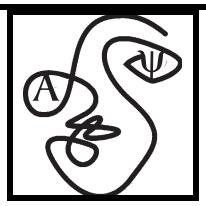
Madden, C. C., Kirkby, R. J., & McDonald, D. (1989). Coping styles of competitive middle distance runners. *International Journal of Sport Psychology*, *20*, 287–296.

Madden, C. C., Kirkby, R. J., McDonald, D., Summers, J. J., Brown, D. F., & King, N. J. (1995). Stressful situations in competitive basketball. *Australian Psychologist*, *30*, 119–124.

Maniar, S. D., Curry, L. A., Sommers-Flanagan, J., & Walsh, J. A. (2001). Student-athlete preferences in seeking help when confronted with sport performance problems. *The Sport Psychologist*, *15*, 205–223.

Prapavessis, H., & Grove, J. R. (1995). Ending batting slumps in baseball: A qualitative investigation. *Australian Journal of Science and Medicine in Sport*, *27*, 14–19.

Taylor, J. (1988). Slumpbusting: A systematic analysis of slumps in sports. (1988). *The Sport Psychologist*, *2*, 39–48.



Personal Initiative and Innovation

Fiona Patterson
City University, London, United Kingdom

1. Slumps Versus Choking and Plateaus
2. Defining Innovation
3. The Innovation Process
4. Person-Level Influences
5. Work Group Innovation
6. Leadership
7. Organizational Innovation
8. Emerging Issues
9. Summary
Further Reading

GLOSSARY

creativity The generation of new and original ideas.
exnovation When an organization actively discards existing practices associated with a previously implemented innovation, thereby allowing the adoption of a new innovation.
incremental innovation An innovation focusing on improvements of existing processes, products, or services to improve performance in an existing market.
innovation The process of generating and applying new ideas to benefit individuals, groups, organizations, or the wider society.
leader–member exchange (LMX) Where leaders develop relationships with all employees in their work groups; these can be high-quality or low-quality relationships in which a pattern of social exchange emerges.
path analysis A set of statistical procedures that allows the researcher to test hypothesized causal relationships among variables (sometimes called causal modeling).
personal initiative A class of active behaviors with three facets: self-starting, proactivity, and persistence.
radical innovation An innovation focusing on the development of processes, products, and services with unprecedented performance potential to create a dramatic change in existing markets.
role innovation Changes in task objectives, responsibilities, methods, materials, and interpersonal relationships.

Innovation is essential for organizational prosperity. It is an iterative process that involves idea generation and idea implementation. There are a variety of influences on the propensity for innovation at work that can be classified into three levels of analysis: individual, work group, and organizational.

1. INTRODUCTION

Over the past two decades, there has been an explosion of interest in innovation as organizations have recognized that creating new processes, products, and procedures is vital for productivity and growth in all sectors. With more dispersed and virtual working, role innovation is essential because clearly defined job descriptions no longer exist for many job roles. The research literature on creativity and innovation is immense, both in magnitude and in diversity. Until recently, there has been confusion in terminology regarding these two concepts, particularly in definitions and in criteria for assessment. Many have observed the lack of a cohesive theoretical

©2004 Elsevier Inc.
All rights reserved.

understanding of how creativity and innovation operate in organizations.

First, the concepts of creativity and innovation are defined. The various factors that influence innovation in organizations can be classified into three main levels of analysis: individual, work group, and organizational. Each of these levels of analysis is reviewed. Finally, a summary of the emerging issues is presented along with some implications for future research and practice.

2. DEFINING INNOVATION

The concepts of creativity and innovation are confused repeatedly, and these terms are often used interchangeably in the literature. Creativity and innovation are overlapping constructs, but the main distinction is with regard to novelty. Creativity is exclusively concerned with generating new and entirely original ideas. Innovation is a broader concept because it also encompasses the application and implementation of new ideas to produce something new and useful (in the context of groups, organizations, or societies). Innovation is often referred to as a process because implementing new ideas necessarily involves influencing others, whereas creativity could be achieved in isolation. Furthermore, an innovation could be the application of something familiar in one organization to another unfamiliar organization.

The most widely accepted definition of innovation is that of West and Farr, emphasizing the positive nature of innovation: "the intentional introduction and application within a role, group, or organization of ideas, processes, products, or procedures, new to the relevant unit of adoption, designed to significantly benefit the individual, the group, the organization, or wider society."

Nearly all authors recognize that innovation in organizations is a complex iterative process. Numerous approaches and process models have been proposed, but two main stages are common to all of them: a suggestion phase and an implementation phase. Innovation is not a linear process given that it involves several cycles of activities such as initiation, reappraisal, adaptation, implementation, and stabilization. In addition to person-level resources, there are various environmental factors that influence this process such as feedback, leadership style, resource availability, and organizational climate. Table I summarizes the factors that have been shown to influence innovation in organizations at the three levels of analysis: individual, work group, and organizational.

3. THE INNOVATION PROCESS

Early research involved stage model approaches to innovation, where authors attempted to identify the sequence of activities during the innovation process. Zaltman and colleagues' stage model of the innovation process in 1973 is perhaps the most widely cited. The model suggests two main phases, initiation and implementation, with various substages in each. The initiation stage involves the substages of knowledge awareness, formation of attitudes, and the decision to adopt or abandon. During the implementation stage, ideas are tested and eventually the innovation is sustained and routinized. Although this research emphasizes that the innovation process is multifaceted, King and Anderson criticized stage-based approaches because they often lack observable supporting evidence. Innovation processes may share some common features, but these are not easily identifiable as discrete stages occurring in a linear fashion.

During the 1990s, Van de Ven and colleagues studied the innovation process in-depth, using empirically grounded observations. Their studies comprised the Minnesota Innovation Research Program (MIRP), captured in detail in *The Innovation Journey* in 1999. This program encompassed several organization-based longitudinal studies of a diverse collection of innovations observed in real time. The research concluded that the innovation process is best characterized as a "nonlinear dynamic system." In this way, the system involves a complex cycle of divergent and convergent activities that may be repeated over time at various organizational levels and that are shaped by forces such as resource supply, leadership behaviors, and the organizational culture. Although the research concluded that innovation is a nonlinear process, by comparing case histories of various innovations, the MIRP studies identified patterns of commonality in innovation processes. These are the initiation, developmental, and implementation periods of the innovation. Each period is described in Table II.

Research demonstrates that different types of innovation may have varying antecedent factors. Furthermore, the evidence suggests that not all of the elements of the innovation process are the same across different innovations. For example, in exploring differences in incremental and radical innovation processes

TABLE I
Factors That Influence Innovation at Work

Level of analysis	*Characteristic*	*Dimension*
Individual	Personality	Openness to experience, low conscientiousness, low agreeableness, originality, imaginative, self-confidence, tolerance of ambiguity, nonconforming, high energy, social rule independent, personal initiative (e.g., self-starting, proactive)
	Motivation	Intrinsic (vs extrinsic), personal initiative (e.g., persistent), determination to succeed, self-efficacy
	Cognitive ability	Above average intellect, creative cognitive style, ideational fluency
	Knowledge	Domain-specific knowledge
	Job characteristics	Autonomy, control, support for innovation, job demands, job dissatisfaction
Work group	Team dynamics and structure	Diversity, conflict, constructive controversy, minority influence, organic structure
	Team climate	Participation, shared vision, support for innovation (articulated and enacted), task orientation
	Team processes	Reflexivity, minority dissent, integration skills, decision-making style
	Leadership style	Transformational, participative, democratic style, delegating responsibility, autonomy, openness to idea proposals and challenges, constructive feedback, expected evaluation, leader–member exchange
Organizational	Structure	Centralization, formalization, complexity
	Climate and culture	Support for risk taking, tolerance of failure, organizational values, reward and recognition for innovation, participative management and decision making, flexible employment contracts
	Size	Number of employees, market share, alliances
	Resources	Economies of scale, slack resources
	Work processes	Collaborative idea flow across functions, idea champions, change agents, suggestion schemes
	Physical environment	Technology to support innovation, informal meeting places, stimulating physical environment

Source. Anderson, De Dreu, and Nijstad (2004).

Note. Factors cited represent a summary of variables, but this is not an exhaustive list. Some factors exert a positive influence, whereas others a negative influence. Some factors have a curvilinear association with innovation, whereas others have an indirect association with innovation.

in 1992, King showed that sequential linear models were most applicable for incremental innovations, whereas radical (revolutionary) innovations were characterized more by convoluted iterative patterns of development.

Innovation processes do not occur in a vacuum and can be shaped by current organizational practices, particularly during the implementation phase. To illustrate this perspective, King and Anderson highlighted the concept of "exnovation." This concept was first introduced in 1981 by Kimberly, who described innovation as a series of processes that combine to define an "innovation life cycle." The final process of the model is exnovation, where the organization must discard existing practice associated with a previously implemented innovation, thereby allowing the adoption of a new innovation, where the life cycle starts again.

4. PERSON-LEVEL INFLUENCES

Innovation involves multiple components at the individual level. However, there has been little synthesis of the literature to build a framework within which to explore the many interrelated characteristics involved. During the past decade or so, there has been a shift in research toward more integrative approaches. For example, in 1999, Sternberg and Lubart proposed an "investment theory" suggesting that the propensity to innovate requires six distinct resources: intellectual

TABLE II
Common Elements in the Innovation Process Identified in the Minnesota Innovation Research Program

Period	*Observation*	*Examples*
Initiation		
Gestation	Initial events might not be intentionally directed at innovation, and gestation might last several years. Coincidental events can trigger a need for change.	Art Fry sang in a church choir for years. Art, like many choir members, put slips of paper in his hymn book to mark each selection. However, his technique was not foolproof, the slips of paper often fell out. Taking his dissatisfaction back to his job, he eventually developed what became 3M's Post-it note pads.
Triggering shocks	Efforts to initiate innovations can be triggered by shocks, whether driven internally or externally to the organization. Without shocks, new ideas might not be acted on.	While on vacation, Edwin Land took pictures of his daughter. When she showed her disappointment that she could not see the results immediately, he set his mind to the task of developing instant photography. In an organization, a shock could be new leadership, a recent product failure, or a change in legislation.
Plans, budgets and exposure	Realistic plans are developed to gain resource commitment (time and finance) to the innovation.	Cost–benefit plans are needed to convince resource controllers and venture capitalists that the risk is worth it. In the organizations studied by the MIRP, during this period, entrepreneurs acted to deflect attention away from the innovations' uncertainty and toward their potential value.
Developmental		
Proliferation	Initial ideas proliferate into diverse multiple pathways, making the process difficult to manage.	In the MIRP studies, most innovations did not consist of a single new product or procedure. Instead, families of related new products were developed in multiple pathways to reduce the risk of failure.
Setbacks	Setbacks can occur frequently because initial plans go wrong or there are unforeseen environmental events.	This period is characterized by budget overruns, unmet deadlines, unsuccessful product launches. Organizations may respond by adjusting plans and resources.
Shifting performance criteria	Criteria for innovation success and failure often change over time, reflecting differing stakeholder views.	Long-term goals could be expressed as return on investment, whereas short-term criteria may be obtaining resources, recruiting appropriate personnel, acquiring equipment, and so on.
Fluid participation of personnel	Personnel often do not work full-time on the innovation. In addition, specific expertise may be required.	There may be high turnover of personnel, or relevant technical expertise might not be available. Part-time involvement means that commitment may vary among personnel.
Investors/Top management involvement	Senior management and/or investors intervene in the process.	The roles of senior management may vary and might not have a unified perspective. Interventions could be to remove impediments or possibly to expand the innovation.
Relationships frequently change	Relationships shape the process, with many interest groups engaging in various transactions to move the innovation forward.	Relationships could be internal to the organization or between external parties through partnerships and joint ventures. With a complex network of interest groups involved, reaching a consensus to move innovations forward might be slow.
Community/ Industry infrastructure	Innovation stakeholders often work within an industry or community infrastructure that is needed to implement an innovation.	The infrastructure could include institutional norms or institutional standards. Organizations seek to establish a distinctive position through innovations but simultaneously must operate within an industrial infrastructure or community.

Continues

Continued

Implementation/Termination		
Integrating the old and the new	The new is linked and integrated with the old rather than substituting or replacing what has gone before.	Because of limited resources and for political reasons, innovations are seldom substituted for existing organizational programs. Instead, the aim is to integrate existing arrangements.
When innovations stop	Innovations terminate when they are institutionalized or when resources expire.	Innovations become routinized and diffused into adopted practice. Reinvention may also occur here when adopters may modify innovations to fit their local circumstances.
Attributions are made	Stakeholders explain causes for the successes or failures of the innovation.	Whether correct or not, senior managers and investors make attributions for success or failure that can significantly affect the career development of the innovation champions and their future likelihood of gaining resources for innovations.

Source. Van de Ven, Polley, Garud, and Venkataraman (1999).
Note. MIRP, Minnesota Innovation Research Program.

abilities, knowledge, styles of thinking, personality, motivation, and environment. This reflects the key areas of research at the person level where previous literature can be classified into that concerning associations between innovation and (1) intelligence, (2) knowledge, (3) personality, (4) motivation, and (5) personal initiative.

4.1. Intelligence

Numerous researchers have explored the association between innovation potential and intelligence. However, the findings are generally inconclusive. Much of the literature in this area can be classified into four categories conceiving of innovation as (1) a subset of general intelligence, (2) an aspect of genius, (3) a set of cognitive abilities and mental processes, and (4) associated with observer judgments of intelligence.

4.1.1. General Intelligence

Early research claimed that creativity was equivalent to high intelligence. The best-known researcher in this field is Guilford. In his theory of the structure of intellect published during the 1950s, he claimed that creative thinking was a mental ability involving divergent production as "thinking that goes off in different directions." Many researchers followed Guilford's work by producing evidence that ideational fluency (i.e., quantity of new ideas) underlies divergent thinking test scores. However, review studies show that such scores often failed to correlate significantly with various indexes of innovation. Other investigations have tested the possibility of a curvilinear relationship between intelligence and innovation where intelligence would potentially become less influential as the level of intelligence increases beyond a certain point. However, tests have proven to be inconclusive, and some authors doubt whether divergent thinking tests measure abilities actually involved in creative thinking.

4.1.2. Aspect of Genius

Some have suggested that genius, as the most obvious manifestation of high intelligence, is closely tied to the propensity for innovation. However, there has been a lack of evidence to support a direct relationship between innovation and intelligence. Many have concluded that intelligence is a necessary, but not sufficient, condition for innovation. Recent studies conclude that intelligence and innovation potential are moderately related but that once intelligence quotient (IQ) scores surpass 115, the relationship is near zero. This finding has been described as "threshold theory," where intelligence and innovation potential, instead of being twin or even sibling constructs, may be more like cousins.

4.1.3. Cognitive Abilities

In 1992, Finke and colleagues suggested that to understand the role of cognitive abilities in idea generation, one must draw on current models in cognitive

psychology and use experimentally based observations of the processes that underlie generative tasks. Their work followed a framework called the "geneplore model." This model proposes that many creative activities can be described in terms of an initial generation of ideas or solutions followed by an extensive exploration of those ideas. Initial ideas are referred to as "preinventive" in the sense that they are incomplete solutions but offer promise in terms of originality and utility. The model assumes that one would alternate between generative and exploratory phases, refining the structures according to the demands or constraints of the specific task. This "creative cognition" approach emphasizes that generative capacity is a property of normative human cognition. Individual differences occur due to variations in the use and application of these generative processes together with the sophistication of an individual's memory and knowledge in the relevant domain. In simple terms, the capacity for creative cognition is normally distributed. Highly creative people do not have minds that operate in any fundamentally different way from those of other individuals.

4.1.4. Observer Judgments of Intelligence

Innovative individuals are often perceived and rated by others as more intelligent than less innovative individuals. For example, in MacKinnon's studies of architects during the 1960s, supervisors rated innovative architects as more "intelligent" than less innovative individuals. MacKinnon described the innovative architects as having high "effective intelligence", and argued that traditional measures of intelligence (e.g., IQ) do not fully explain this "real-world" intelligence. In 2003, Feist and Barron showed that observer-rated intelligence at 27 years of age predicted lifetime innovation at 72 years of age. Similarly, tested intelligence had much weaker relationships with innovation over this time.

Historically, the literature on innovation and intelligence has lacked clarity. Part of the problem has been that intelligence, similar to innovation, is often viewed as a unitary concept. Previous theories of intelligence have tended to overemphasize cognitive abilities and downplay the role of knowledge-based intelligence.

4.2. Knowledge

Nearly all researchers in this field have assumed that knowledge is a key variable in both generative thinking and innovation. Immersion in domain-specific knowledge is an essential prerequisite for innovation because an individual must have an accurate sense of domain (i.e., contextual relevance) before he or she can hope to change it for the better. However, the literature highlights that too much expertise in one area can also be a block to innovation within that domain. There is an "inverted U" relationship between knowledge and innovation, where too much or too little knowledge will not lead to new inventions. During the 1980s, Simonton studied the lives of more than 300 eminent people to explore life span development of innovation and found that both a lack of familiarity and an excess of familiarity within a subject domain can be detrimental to innovation. Domain knowledge, like intelligence, is necessary but not sufficient for innovation to occur.

4.3. Personality

From several decades of research on the association between innovation and personality, a consistent set of characteristics has emerged. These include imaginative, inquisitive, high energy, high desire for autonomy, social rule independence, and high self-confidence. The Five Factor Model (FFM) of personality has become a nearly universal template with which to understand the structure of personality. The FFM dimensions are Openness to Experience (e.g., ideas, aesthetics), Agreeableness (e.g., compliance, straightforwardness), Conscientiousness (e.g., order, dutifulness, competence), Extraversion (e.g., warmth, gregariousness, activity), and Neuroticism (e.g., anxiety, depression). Given that the FFM is an appropriate model for charting individual differences among adult populations, it provides a useful structure to review the literature exploring associations between personality and innovation.

4.3.1. Openness to Experience

There is good empirical evidence of a positive association between various characteristics associated with innovation and those used to depict openness (e.g., imaginative, original, flexible, unconventional). Research suggests that openness enhances an individual's intrinsic motivation toward novelty and, therefore, works in a multiplicative way to produce innovation. Although there are some inconsistencies in the findings, openness is perhaps the most important personality dimension to predict the propensity for innovation.

4.3.2. Agreeableness

Several studies have demonstrated a negative association between agreeableness and innovation. Innovators have high social rule independence. During the 1990s, Eysenck emphasized the potentially negative characteristics of innovators, where innovators are often outspoken, uninhibited, quarrelsome, and/or asocial.

4.3.3. Conscientiousness

Research has demonstrated that lack of conscientiousness is associated with innovation. Defined by terms such as fastidious, ordered, neat, and methodical, the evidence shows that individuals high on conscientiousness are more resistant to changes at work and are more likely to comply with current organizational norms. Recent studies have shown that the facets of conscientiousness most closely associated with innovation are being methodical, being ordered, and being dutiful.

4.3.4. Extraversion

Meta-analytic studies (i.e., large-scale reviews of quantitative studies) suggest that, in general, introversion is positively associated with innovation. Similarly, many have argued that isolation and withdrawal are necessary conditions for generating new ideas. However, there is little evidence from organizational contexts, and some have suggested that extraversion is a positive predictor of innovation. In meta-analytic studies of occupational work performance in general, extraversion has been shown to be a positive predictor for many occupations. This is particularly the case in large organizations where interpersonal factors are likely to be important for effective job performance (e.g., sales, managers, other professional occupations). Further research is needed to explore the association between extraversion and innovation.

4.3.5. Neuroticism

There is relatively little research examining an association between neuroticism (i.e., low emotional stability) and innovation. Of what literature is available, there appear to be some inconsistencies depending on the domain of interest. For example, in a meta-analytic review, Feist observed that artists appear to be more anxious, emotionally labile, and impulsive than the scientists he studied. A more thorough investigation in this area is necessary, particularly in organizational settings with a broader range of occupations. Some have suggested a curvilinear association between emotional stability and performance, where too much or too little anxiety is detrimental to innovation and where moderate levels of anxiety can enhance innovative potential.

There are various personality measures available that claim to assess the propensity to innovate, including Cattell and Eber's Sixteen Personality Factor (16-PF) Questionnaire and Patterson's Innovation Potential Indicator. Other approaches include cognitive-style instruments such as Kirton's Adaption–Innovation Inventory.

4.4. Motivation

High levels of motivation are required for innovation, and innovators are viewed as displaying a devotion and total absorption in their work. During the 1980s, Amabile suggested a componential model of innovation that involves three components: intrinsic task motivation, domain-relevant skills (i.e., expertise), and innovation-relevant process skills (i.e., cognitive skills and work styles conducive to novelty). The model includes a five-stage description of the innovation process—task presentation, preparation, idea generation, idea validation, and outcome assessment—where the roles of the three components vary at each of the stages. Amabile's model suggests how and where individual skills and motivation affect the progress of the innovation process.

Intrinsic motivation is clearly a prerequisite for innovation, but the role of extrinsic motivators is less clear. In exploring environmental influences on motivation, the evidence suggests that constructive evaluation (e.g., informative, supportive, recognizes accomplishment) can enhance innovation. Amabile suggested that intrinsic motivation is particularly important in tasks that require novelty and that extrinsic motivators may be a distraction during the early stages of the innovation process. Later during the innovation process, when persistence and evaluation of ideas are required, synergistic extrinsic motivators may help innovators to persist in solving the problem within the domain. Any extrinsic motivator that enhances an employee's sense of competence without undermining self-determination should enhance intrinsic motivation and, thus, increase the propensity for innovation.

In related studies, results showed that individuals who received positive feedback given in an informational style, and who worked in a high-task autonomy environment, generated the most innovative solutions. Perceived competence (i.e., an individual's belief that

he or she is capable of doing the task) and self-determination (i.e., an individual's perception that he or she is the main influence behind successful performance of the task) are the two main antecedents of intrinsic motivation. When autonomy is high (in choosing work methods and work schedules) and when an individual receives positive informational feedback, this is likely to optimize innovative production.

4.5. Personal Initiative

Based on work by Frese and colleagues, the concept of personal initiative (PI) describes a class of behaviors that have been positively linked with innovation and entrepreneurial orientation. PI is defined by three main facets: self-starting, proactivity, and persistence. A self-starting approach is characterized by setting oneself context-specific goals and going beyond formal job requirements. Proactivity implies that one anticipates opportunities and problems rather than reacts and prepares to deal with them before they occur. PI also involves persistence because this will be required to overcome barriers in reaching one's goals. Frese suggested that these three facets of PI reinforce each other and tend to co-occur. Taking an action-based approach, each of the three facets of PI can be used to understand how individuals develop goals, collect information, and make plans for executing them as well as how individuals gather and use feedback.

In 2001, Frese and Fay proposed a complex model of distal and proximal factors that influence performance in organizations, whereby environmental supports (e.g., support for PI, control at work), knowledge, skills and abilities (e.g., cognitive ability), and personality influence individual orientations (e.g., self-efficacy, handling errors, active coping), which in turn influence PI. In this way, PI is conceptualized as a set of active behaviors and is directly linked to effective performance in organizations.

There is a growing body of research examining the association between PI and innovation. The relationship between PI and conservatism in the workplace has been studied. Results show that conservatives showed significantly less PI, were less orientated toward growth and challenge, and were less innovative compared with their less conservative counterparts. In exploring environmental influences on PI, research shows that increases in control and job complexity can help to enhance PI at work.

PI is particularly important during the idea implementation phase of the innovation process because it involves overcoming barriers and persistence. In a study using path analysis, several predictors were examined, including individual-level variables (e.g., PI, self-efficacy, interest in innovation), work characteristics (e.g., control, complexity), motives (e.g., reward), and organizational system factors (e.g., supervisor support). Results showed that being proactive, being actively involved in one's work environment, and being confident that one is capable of thinking of good ideas were the most important predictors of innovation. Organizational factors were also important variables, and the results suggest that innovation is maximized when organizational climates promote an active approach toward work and interpersonal risk taking.

The concept of PI has been applied at an organizational level. Findings suggest that companies with pro-initiative climates are more innovative and profitable. Climates for initiative may improve organizations' ability to deal with innovation and change by encouraging self-starting, proactivity, and persistence in employees.

5. WORK GROUP INNOVATION

Until recently, there had been little literature available regarding innovation in work groups. Several factors enhance group innovation, and four key issues are considered: (1) operating principles in groups; (2) diversity, dissent, and minority group influence; (3) group development over time; and (4) group climate.

5.1. Operating Principles

In 2002, King and Anderson highlighted the differences between "organic" and "mechanistic" group structures in their discussion of operating principles. An organic group structure is characterized by loose boundaries of authority and responsibility and by a propensity to work as a group rather than breaking projects down into discrete tasks for individuals. Conversely, mechanistic groups are typified by being rule bound, hierarchical, and formal in operation. Organic group structures tend to be more innovative because autonomy and freedom are enhanced. Work group autonomy is an essential ingredient of innovation, and it is most effective when combined with unambiguous goals and objectives from management. Autonomy is important for idea generation, and mechanistic forms of organizing may have a role in coordinating the implementation of ideas. Research

suggests that "semi-autonomous" operating groups are optimal for innovation.

During the 1990s, West and colleagues described the process of "team reflexivity" as the extent to which team members collectively reflect on the team's objectives, strategies, and processes, as well as the wider organization context, and adapt them accordingly. When teams regularly devote time to collective consideration of their objectives, plans, and actions, they are in a position to develop continuously and potentially to improve performance. Reflexivity enhances team learning and allows teams to refine shared goals, plans, and actions, all of which are likely to enhance the likelihood of innovation.

5.2. Diversity, Dissent, and Minority Group Influence

Groups consisting of people with a wide variety of backgrounds and perspectives are more likely to consider a wider variety of approaches to tasks. However, too much diversity will increase the likelihood of conflict within the group, and this could have negative consequences for productivity. During the 1980s, Tjosvold introduced the concept of "constructive controversy," which indicates the value of social interaction, and specifically controversy, in decision making. Constructive controversy occurs when team members believe that they are in a cooperative environment (emphasizing mutually beneficial goals rather than a "win–lose" solution), when they believe that their personal competence is acknowledged, and when members use processes of mutual influence rather than attempted dominance. Accepting and encouraging the expression of minority views and dissent in groups and organizations is an important stimulus for creativity and innovation. Conversion to a minority view in groups is most likely to occur when the minority is consistent and confident in presenting arguments. A behavioral style of persistence is most likely to promote attitude change and influence over the majority. The minority group must reflect a visible commitment to the norms, values, and interests of the majority to gain influence.

5.3. Group Development over Time

In general, the longer groups are together, the less innovative they become as teams grow to be habit bound and more resistant to change. Work groups do not come together and develop over time in a vacuum. Initiating change in an organization might not be possible if this occurs at a peak time in an organization's annual cycle of activity. The concept of "entrainment" has been used to explain that cycles of activities are paced by the numerous other cycles. In applying this to organizations, this "entraining process" provides windows (or barriers) of opportunity for change to take shape and take root. In addition to longevity, the timing of the formation of a project group can be critical to its success.

5.4. Group Climate

During the 1990s, West and Anderson proposed a four-factor model of team climate. This model suggests that group innovation is related to (a) participative safety (members feel psychologically safe in proposing new and improved ways of doing things, and all participate in decision making), (b) vision (the team's goals and objectives are clearly defined, shared, attainable, and valued), (c) support for innovation (there is expectation, approval, and practical support toward group members' attempts to introduce new and improved ways of doing things in the work environment), and (d) task orientation (the team is committed to achieving the highest possible standard of task performance using constructive progress-monitoring procedures).

There is good evidence to support the existence of these four factors in relation to group climate. Groups with clearly defined and shared visions are more likely to develop new working methods because their efforts are focused and have direction. Participation in decision making is important because it increases the likelihood that group members commit to decision outcomes and will be willing to offer new ideas. In addition, innovation requires group commitment to achieve high task performance and requires members to offer articulated and enacted support for innovation attempts.

6. LEADERSHIP

Since the early 1980s, several independent studies have identified a range of leadership behaviors that enhance employee innovation. These behaviors include encouragement of risk taking, an open style of communication, a participative and collaborative style, provision of autonomy and freedom, support for innovation (verbal and enacted), constructive feedback, and optimism about the future. Most research has focused on one or two management behaviors resulting in a lack of integration of results. Furthermore, because many studies have focused solely on either idea generation or the implementation of ideas, clarity has been difficult to achieve.

King and Anderson proposed a contingency model to understand the role of leadership in relation to four phases of the innovation process: (a) initiation (the leader encourages ideas and is nurturing, supportive, open-minded, and nonjudgmental), (b) discussion (the leader obtains opinions, evaluates, and agrees on plans), (c) implementation (the leader sells the solution to all stakeholders, designs plans to include stakeholders, and gains commitment to plans), and (d) routinization (the leader checks effectiveness, modifies, and refines). Therefore, leaders must be flexible in their approaches and must employ different styles according to the different phases of the innovation process.

The quality of the leader–member exchange (LMX) predicts the propensity for innovation. Leaders develop relationships with all members in their work groups, and these can be characterized as high-quality or low-quality relationships, where a pattern of social exchange emerges. In 1999, Tierney and colleagues demonstrated that a high-quality LMX predicts employee innovation because employees who were engaged in more challenging tasks received more resources and rewards for innovation. In recent studies, job self-efficacy has been shown to be the strongest predictor of creative self-efficacy. This implies that if managers are to enhance employee innovation, they must provide the necessary training and feedback to enhance employee confidence and mastery. Recent theories of leadership describe two styles: a transformational style (inspiring and motivating others to achieve a shared vision) and a transactional style (influencing others through contingent rewards and punishments). Research demonstrates that a transformational leadership style enhances employee and group innovation.

7. ORGANIZATIONAL INNOVATION

Several factors facilitate or inhibit innovation at the organizational level of analysis: (1) climate and culture, (2) structure and size, (3) resources and work processes, and (4) external demands and environment.

7.1. Climate and Culture

Climate and culture are important antecedents to innovation to the extent that risk taking and idea generation are encouraged. Climate is the shared perceptions and attitudes that influence the social interaction behavior. Culture is a set of commonly shared beliefs, values, and assumptions that form the basis for climate behavior. Supportive and challenging climates encourage innovation, whereas environments characterized by distrust, personal hostilities, limited autonomy, and unclear work goals inhibit the implementation of ideas. Organizational support for innovation is characterized by support for members pursuing new ideas and encouragement of innovation through both words and deeds. The espoused expectations and organizational values in relation to innovation (e.g., openness to change, willingness to experiment with ideas) are likely to have a major influence on employee behavior.

Amabile and colleagues noted that organizational encouragement is a vital aspect of a work environment for innovation and includes (a) encouragement of risk taking and a valuing of innovation from the highest level, (b) fair and supportive evaluation of ideas, (c) reward and recognition for innovation, (d) collaborative idea flow across the organization, and (d) participative management and decision making.

Research shows that an organizational climate can build a perceived need for innovation. Setting innovation goals at both the team and individual levels is one way in which to signal that innovation is valued and innovation can be incorporated in performance management systems. In 2001, Csikszentmihalyi argued that innovation must be considered within a complex system composed of the domain, field, and person. Thus, organizations have gatekeepers who can make decisions about which innovations are worth supporting without being too accepting or too conservative with regard to novelty. Organizational practices that positively influence innovation include flexible employment contracts and knowledge management systems, whereas practices that seek to control employees lower innovation supportive cultures.

7.2. Structure and Size

There are three main aspects regarding the influence of organizational structure on innovation: centralization, formalization, and complexity. Centralization refers to the extent to which authority and decision making lie at the top of a hierarchy. Centralization is believed to hinder innovation because it restricts information flow and communication. Decentralization also gives rise to greater participation by allowing more viewpoints to be considered during idea generation. However, decentralization has a negative impact on the adoption of innovation and idea implementation in that

decentralization is necessary for idea generation, whereas centralization is important for idea implementation.

Formalization is the degree of emphasis placed on rules and procedures. High formalization inhibits innovation because it prevents organizational decision makers from seeking new sources of information and from making changes at work. However, procedural guidelines and regulations can enhance the successful implementation of new ideas.

Complexity is the degree of occupational specialization in the organization and is believed to enhance idea generation and inhibit idea implementation. High complexity is likely to provide a diverse range of information sources during idea generation, but such diversity could accelerate the potential for conflict during the implementation phase.

Because innovation is a multifaceted process, some variables that are important to enhance idea generation may be counter to promoting idea implementation.

Some claim that increased organizational size enhances innovation because there is greater access to resources through economies of scale. However, organizational size does not directly predict innovation because it is related to several different variables such as resource availability (which promotes innovation) and integration of communication and knowledge sharing (work processes often become more fragmented as size increases and so innovation is hindered).

7.3. Resources and Work Processes

Availability of resources is positively related to innovation because resources are needed to develop ideas. Beyond a certain threshold, more available resources do not significantly increase the likelihood of innovation. Researchers have observed a concept of "escalation"; that is, there may come a point when there is so much invested that it is too costly to stop the implementation process.

Integration across functions and groups promotes innovation through establishing common goals, knowledge sharing, and communication. For example, mergers have been shown to reduce the potential for innovation because resources are focused on managing structural changes to enhance amalgamation. Conversely, alliances and joint ventures between competitors or between customers and suppliers promote innovation and product development where combining resources and knowledge is beneficial.

Some organizations create "idea champions" (i.e., formal job roles to enhance innovation) or "change agents" (whose specific responsibility is to encourage and enhance innovation at work). Some organizations create an innovation resource fund specifically to encourage development and implementation of new ideas. Idea capture or suggestion schemes are beneficial to enhancing employee involvement in generating and implementing new ideas.

A stimulating physical environment can influence work processes to enhance innovation. Some organizations have innovation laboratories that make use of technologies such as electronic brainstorming software. Using this software in groups can significantly enhance the quantity of ideas produced because it reduces some of the negative social factors in paper-based group brainstorming such as evaluation apprehension (where members withhold ideas because they fear negative evaluation) and free riding (where members rely on others to accomplish the task).

7.4. External Demands and Environment

The context in which organizations operate influences the propensity for innovation. For example, environmental uncertainty predicts administrative innovation, and low market share predicts higher levels of product innovations in some sectors. In this way, uncertainty generated by external demands and potential threats can stimulate innovation. However, under extremely threatening conditions, some organizations tend to centralize control, conserve resources, restrict information flow, and rely on traditional work methods. As a result, the propensity for innovation may be reduced in the long term.

The information flow in and out of organizations will influence innovation to the degree that there is an active search of the environment for new ideas. The external reputations of organizations can influence innovation because innovative organizations may be more likely to attract individuals with a high innovation potential.

8. EMERGING ISSUES

Innovation is a process with a variety of situational- and individual-level determinants Each determinant may be more or less important during the idea generation or idea implementation phase. Similarly, the role of the environment in shaping innovation is likely to have a differential impact on individuals. Future research should explore interaction effects where the

individual differences in personality or experience may influence the way in which a person interprets or reacts to extrinsic rewards or constraints. In general, there is a lack of theoretical integration of the variables involved at each level of analysis and relevant interaction effects. In 2004, Anderson and colleagues reviewed previous innovation research and argued for the development of multilevel theories and study designs and for the triangulation of research methods to address these issues. The need to measure negative predictors of innovation, such as job dissatisfaction, task conflict, and home/work strain, was emphasized. Future research could explore innovation as the independent variable.

Historically, there has been less research targeting group-level innovation, but this has expanded recently and there are now clear practical recommendations regarding enhancement of innovation in teams. Future studies on innovation should reflect the changing nature of work and the use of technology. The impact of dispersed and virtual teams (where members are geographically dispersed and in different global time zones) on team dynamics and innovation has not been explored. Similarly, the impact of international and cultural differences on team working and innovation (where team selection may also be important) has yet to be explored. Future research could examine the impact of technology and the physical environment on innovation by evaluating innovation laboratories to enhance idea generation and implementation at work.

The role of leaders in enhancing innovation has been neglected. Leaders are the gatekeepers in the innovation process and have a significant impact on the organizational climate and culture. Although organizations recognize the need for innovation, some managers may be less equipped to recruit and develop innovative employees.

9. SUMMARY

Innovation is a multifaceted and multilevel process. Figure 1 summarizes the main elements of the innovation cycle and the various influences at different levels

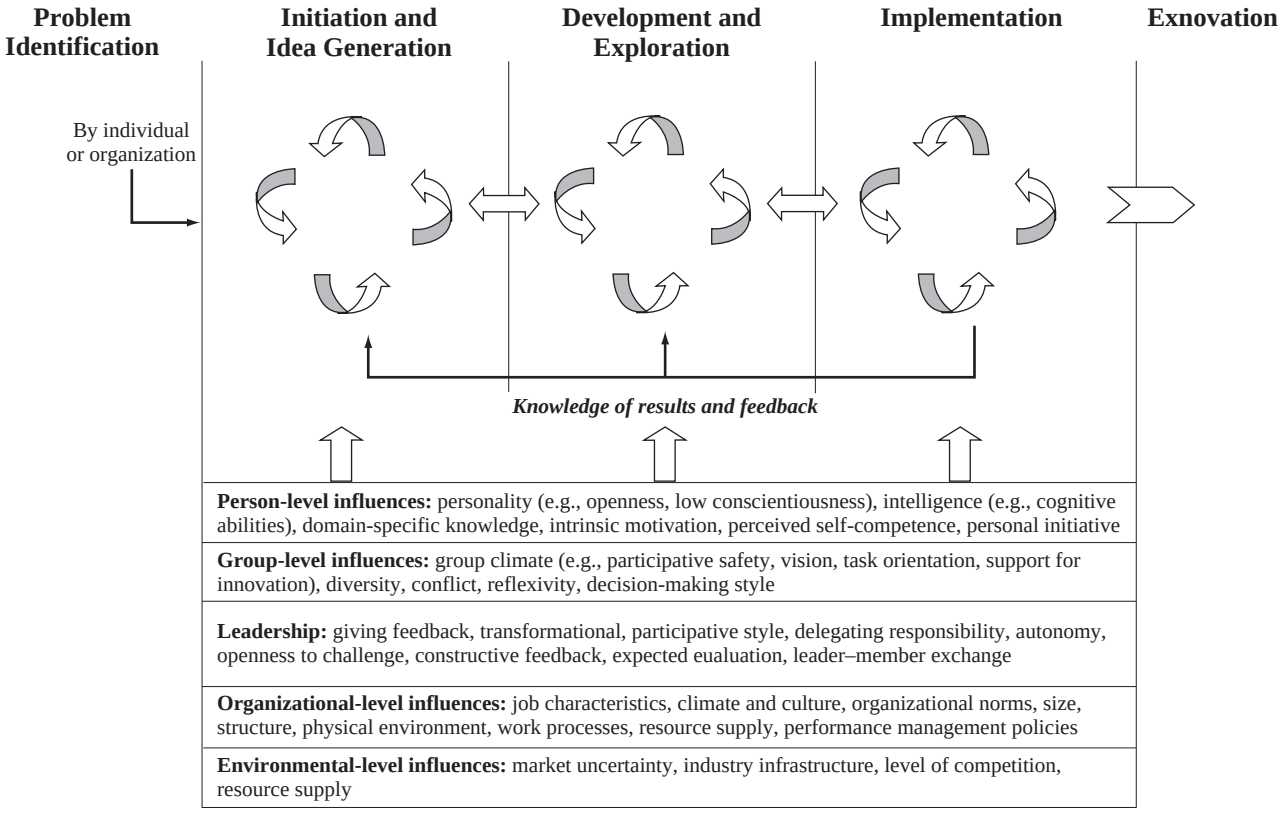

Problem Identification
Initiation and Idea Generation
Development and Exploration
Implementation
Exnovation
By individual or organization
Knowledge of results and feedback
Person-level influences: personality (e.g., openness, low conscientiousness), intelligence (e.g., cognitive abilities), domain-specific knowledge, intrinsic motivation, perceived self-competence, personal initiative
Group-level influences: group climate (e.g., participative safety, vision, task orientation, support for innovation), diversity, conflict, reflexivity, decision-making style
Leadership: giving feedback, transformational, participative style, delegating responsibility, autonomy, openness to challenge, constructive feedback, expected eualuation, leader–member exchange
Organizational-level influences: job characteristics, climate and culture, organizational norms, size, structure, physical environment, work processes, resource supply, performance management policies
Environmental-level influences: market uncertainty, industry infrastructure, level of competition, resource supply

FIGURE 1 Phases in the innovation process and influencing variables.

of analysis. Innovative individuals have a creative cognitive style and are open to new experiences. Innovators are proactive, knowledgeable about the domains in which they are trying to implement their ideas, and persistent in the face of obstacles. A high level of diversity among employees enhances team innovation if there is integration of ideas among employees and team members. A positive team climate can have a significant influence on innovation and is characterized by participative safety, vision, support for innovation, and task orientation. Leaders can enhance innovation and may play different roles during different phases of the innovation process (e.g., encouraging ideas during idea generation, providing adequate resources and selling ideas to others during idea implementation). Many organizational factors influence innovation, including informal characteristics (e.g., climate and culture), formal policies (e.g., performance management systems), and factors external to the organization (e.g., market demand and uncertainty).

See Also the Following Articles

Agreeableness ■ Intelligence and Culture ■ Intergroup Relations and Culture ■ Leadership and Culture ■ Organizational Culture and Climate ■ Traits ■ Work Motivation ■ Work Teams

Further Reading

Amabile, T. M. (1988). A model of creativity and innovation in organizations. *Research in Organizational Behavior*, *10*, 123–167.

Amabile, T. M. (1996). *Creativity in context: Update to the social psychology of creativity*. Boulder, CO: Westview.

Anderson, N., De Dreu, C. K. W., & Nijstad, B. A. (2004). The routinization of innovation research: A constructively critical review of the state-of-the-science. *Journal of Organizational Behavior*, *25*, 147–173.

Cattell, R. B., & Eber, H. (1970). *Handbook for the Sixteen Personality Questionnaire*. Champaign, IL: Institute for Personality and Ability Testing.

Csikszentmihalyi, M. (2001). The context of creativity. In W. Bennis, & G. Spreitzer (Eds.), *The future of leadership: Today's top leadership thinkers speak to tomorrow's leaders* (pp. 116–124). San Francisco: Jossey–Bass/Pfeiffer.

Eysenck, H. (1995). *Genius: The natural history of creativity*. Cambridge, UK: Cambridge University Press.

Feist, G. J., & Barron, F. X. (2003). Predicting creativity from early to late adulthood: Intellect, potential, and personality. *Journal of Research in Personality*, *37*, 62–88.

Finke, R. A., Ward, T. B., & Smith, S. M. (1992). *Creative cognition: Theory, research, and applications*. Cambridge, MA: MIT Press.

Frese, M., & Fay, D. (2001). Personal initiative: An active performance concept for work in the 21st century. *Research in Organizational Behavior*, *23*, 133–187.

King, N. (1992). Modelling the innovation process: An empirical comparison of approaches. *Journal of Occupational and Organizational Psychology*, *65*, 89–100.

King, N., & Anderson, N. (2002). *Managing innovation and change: A critical guide for organizations* (2nd ed.). London: Thomson.

Kirton, M. (1976). Adaptors and innovators: A description and measure. *Journal of Applied Psychology*, *61*, 622–629.

Patterson, F. (2000). *The Innovation Potential Indicator: Test manual and user's guide*. Oxford, UK: OPP.

Patterson, F. (2002). Great minds don't think alike? Person level predictors of innovation at work. *International Review of Industrial and Organizational Psychology*, *17*, 115–144.

Simonton, D. K. (1991). Emergence and realization of genius: The lives and works of 120 classical composers. *Journal of Personality and Social Psychology*, *61*, 829–840.

Sternberg, R. J. (Ed.). (1999). *Handbook of creativity*. Cambridge, UK: Cambridge University Press.

Tierney, P., Farmer, S. M., & Graen, G. B. (1999). An examination of leadership and employee creativity: The relevance of traits and characteristics. *Personnel Psychology*, *52*, 591–620.

Tjosvold, D. (1998). The cooperative and competitive goal approach to conflict: Accomplishments and challenges. *Applied Psychology: An International Review*, *47*, 285–313.

Van de Ven, A. H., Polley, D. E., Garud, R., & Venkataraman, S. (1999). *The innovation journey*. New York: Oxford University Press.

West, M. A. (2001). The human team: Basic motivations and innovations. In N. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswesvaran (Eds.), *Handbook of industrial, work, and organizational psychology*, Vol. 2: *Organizational psychology* (Vol. 2, pp. 270–288). London: Sage.

West, M. A., & Anderson, N. A. (1998). Measuring climate for work group innovation: Development and validation of team climate inventory. *Journal of Organizational Behavior*, *19*, 235–258.

West, M. A., & Farr, J. L. (Eds.). (1990). *Innovation and creativity at work: Psychological and organizational strategies*. Chichester, UK: Wiley.

Zaltman, G., Duncan, R., & Holbek, J. (1973). *Innovations and organizations*. New York: John Wiley.